STRUCTURAL CHANGE AND UNIT ROOTS

BY

IN-MOO KIM

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1991

To my parents, my wife and my son

ACKNOWLEDGEMENTS

My hearty thanks are due to my supervisor, Professor G. S. Maddala, for his encouragement, suggestions, and keen intuition. His consolation for my grief over the loss of my mother gave strength to me throughout the completion of my dissertation.

I would also like to thank Professors Mark Rush, Prakash Loungani, and J. S. Shonkwiler for their kind help. My working experience with Professor Prakash Loungani was especially helpful.

The education and guidance I received from Professor Hyun Sik Chung, my ex-supervisor in Korea, was my inspiration during my study in Florida. I am sincerely grateful to him.

I realize that this dissertation involved the invaluable support and sacrifices of my family members, especially my mother and my wife, Jong Han. With thanks from my whole heart, I dedicate this dissertation to them.

TABLE OF CONTENTS

Abstract of Dissertation Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirement for the Degree of Doctor of Philosophy

STRUCTURAL CHANGE AND UNIT ROOTS

By

In-Moo Kim

May 1991

Chairman: Dr. G. S. Maddala
Major Department: Economics

Detection of a structural change in a regression model has recently attracted considerable interest in the literature on both econometrics and statistics. Perron's hypothesis, that standard tests of the unit root hypothesis against trend stationary alternatives cannot reject the unit root hypothesis if the series has a structural break at some intermediate date, has initiated a series of testing procedures for parameter constancy in nonstationary time series. After Christiano's criticism about Perron's a priori known break point, the empirical results from several testing methods endogenizing the break point selection procedure have provided evidence against Perron's hypothesis.

Bayesian inferential procedures for detecting a structural break in dynamic models are more manageable than those of the classical approach. Also, the Bayesian inferential theory is largely unaffected by the presence of unit roots. In this dissertation I review the Bayesian inferential procedure for detecting a structural change in autoregressive models with a Monte Carlo study and apply this methodology to the GNP series of the OECD countries.

Monte Carlo studies about detecting a structural change in the autoregressive model show that the Bayesian posterior mass function of m detects a break point more readily than the classical approach even when the series is nonstationary. When a peak of the marginal posterior mass function of m occurs within a sample period, it indicates a break point.

Using Bayesian inference, I found strong evidence supporting Perron's hypothesis even after endogenizing the break point selection procedure. The results for Canada, France, Italy, and Japan show that standard tests of the unit root hypothesis were biased in favor of accepting the unit root hypothesis because of a structural break.

CHAPTER 1

INTRODUCTION

Detection of a structural change in a regression model has recently attracted considerable interest in the literature on both econometrics and statistics. The conventional statistical tests for parameter constancy in stationary time series suffer from the lack of an appropriate asymptotic theory when a structural break is assumed to be unknown. Recently the functional central limit theorem and the continuous mapping theorem have been used to shed light on the asymptotic theory for testing parameter constancy.

Perron's (1989) hypothesis, that standard tests of the unit root hypothesis against trend stationary alternatives cannot reject the unit root hypothesis if the series has a structural break at some intermediate date, has initiated a series of testing procedures for parameter constancy in nonstationary time series. After Christiano's (1988) criticism about Perron's a priori known break point, the empirical results from several testing methods endogenizing the break point selection procedure have provided evidence against Perron's hypothesis.

Classical testing procedures for parameter constancy require the sophisticated concepts of Brownian motion to obtain an asymptotic theory of testing procedures, especially for the nonstationary case. Tabulating the critical values of test statistics is also extremely difficult since the test statistics are usually functions of Wiener processes.

Bayesian inferential procedures for detecting a structural break in dynamic models are more manageable than those of the classical approach. Also, the Bayesian inferential theory is largely unaffected by the presence of unit roots. In this dissertation I review the Bayesian inferential procedure for detecting a structural change in autoregressive models with a Monte Carlo study and apply this methodology to the data sets analyzed by Banerjee, Dolado, and Galbraith (1990), and Banerjee, Lumsdaine, and Stock (1990).

The flat-prior Bayesian approach has been criticized by Phillips (1990). Phillips suggested an alternative ignorance-prior based on Jeffreys' theory of invariance instead of the flat-prior. I discuss this criticism and report some Monte Carlo results, arguing that Phillips' ignorance-prior is not better than the flat-prior.

Chapter 2 reviews the classical testing procedures for parameter constancy in stationary and nonstationary cases. Chapter 3 derives the Bayesian inferential procedure for

detecting a structural change in autoregressive models and reports some Monte Carlo evidence for identification of a break point. Chapter 4 reviews the previous results analyzed by the classical approach. Chapter 5 reports the empirical results reanalyzed by the Bayesian methodology derived in Chapter 3 for the data sets in Chapter 4. Chapter 6 discusses the criticism of the flat-prior Bayesian inference and the classical approach in the presence of unit roots. Chapter 7 concludes the dissertation.

CHAPTER 2

TESTS FOR PARAMETER CONSTANCY

This chapter describes and compares different types of classical techniques for testing the null hypothesis of constant parameters over time when regression analysis is applied to time-series data. First I will discuss the testing procedures for parameter constancy in the regression model with stationary regressors in Section 2.1. I will briefly sketch how to get an asymptotic theory and review conventional tests for parameter constancy against the different characterizations of alternative hypotheses -- a single structural change or general unspecified alternatives. In Section 2.2 I will sketch two different approaches to obtaining a nuisance parameter free limiting distribution and then discuss several testing procedures for parameter constancy in nonstationary time series.

2.1. Stationary Regressors

The conventional tests for parameter constancy can be classified into two different approaches according to the

specification of an alternative hypothesis. One approach is to test the null hypothesis of constant coefficients against the alternative that a single structural change has occurred at some unknown time. The other is to test against general unspecified structural changes such as varying parameters and random walk coefficients.

Before reviewing the tests, I will sketch briefly how to get an asymptotic distribution which is the essential part of the most recent papers concerning the testing procedures for structural changes.

2.1.1. Brief Sketch of Asymptotic Theory

Recent developments in functional central limit theory (FCLT) and the continuous mapping theorem (CMT) have provided the missing tools to develop an asymptotic theory. The following procedure to get an asymptotic theory can be found in most recent studies about testing structural changes.

1) First express the sequence of a test statistic as a function of $r = m/T$ where m is a split point. The sequence is viewed as a function of r for values in K, some compact subset of (0,1).

2) The partial sums of variables in the sequence of a test statistic can be substituted asymptotically by

Wiener process, $W(r)$, based on the functional central limit theorem. The sequence of a test statistic will be the function of Wiener processes. Note that applying FCLT holds only for change points evaluated in some closed region on (0,1). This means that change points too close to the sample boundaries cannot be tested[1].

3) The test statistic will be some function of the sequence, for example,

$f(\cdot) = \max$: for detecting a single change point

$f(\cdot) = mean$: for general unspecified alternatives.

The asymptotic distribution of the test statistic can be obtained by the continuous mapping theorem. The asymptotic distribution is usually a function of Brownian bridges, $W^o(r)$.

4) The hitting probability of Brownian bridge is obtained by well-known distribution function.

$$Pr[\sup_{0 \le r \le 1} \|B(r)\|_\infty \le X] = 0, \quad x < 0$$

$$= [1 + 2\sum_{i=1}^{\infty} (-1)^i \exp(-2i^2x^2)]^k, \quad x \ge 0$$

This arises in various contexts. For instance, it equals the null distribution of the Kolmogorov-Smirnoff test for k=1. (Billingsley, 1968). Alternatively the

---

[1] Andrews (1990) points out that the sequence of the test statistics at points bounded away from the beginning and end of the sample is an important component of the asymptotic theory. Andrews suggested the practical rule of taking K=[.15, .85] which can be compared with [.10, .90] by Kim and Siegmund (1989).

limiting distribution can be approximated by Laplace transformation or some expansion methods.

2.1.2. Single structural change

The traditional Chow (1960) test is developed to test the null hypothesis of parameter constancy against the alternative of a known break point a priori under the assumption of constant variances. The papers by Quandt (1958, 1960) discuss testing the null hypothesis of constant coefficients against a more general alternative, where a structural change has occurred at some unknown time and the error variance is also allowed to change. However, Quandt's likelihood ratio testing procedure suffers from the lack of a distribution theory. Recently the functional central limit theorem and the continuous mapping theorem have been used to shed light on the asymptotic distribution theory of that. Here I will review Quandt's LR test and then the Kim and Siegmund (1989) and Chu (1989) testing procedures.

2.1.2.1. Quandt's LR Test

In this case the observations are thought, for theoretical reasons, to have been generated by two distinct

regression regimes. Thus, for some subset I of the n observations,

$$Y_i = X_i\beta_i + u_{1i} \qquad (i \in I)$$

and for the complementary subset J

$$Y_j = X_j\beta_j + u_{2j} \qquad (j \in J)$$

The essence of this simple formulation is that all the observations up to the unknown time m come from one regime and all the observations after that point come from the other. On the assumption that m is the time at which the switch from one regime to the other occurred, the likelihood of the sample can be written as

$$L = \left(\frac{1}{\sqrt{2\pi}\sigma_1}\right)^t \left(\frac{1}{\sqrt{2\pi}\sigma_2}\right)^{T-t} \exp\left[-\frac{1}{2\sigma_1^2}(Y_t - X_t\beta_1)'(Y_t - X_t\beta_1) - \frac{1}{2\sigma_2^2}(Y_{T-t} - X_{T-t}\beta_2)'(Y_{T-t} - X_{T-t}\beta_2)\right]$$

The value of L may be evaluated for all possible choices of t and that value chosen as the estimate of the unknown switching point which maximizes L. If we wish to test the hypothesis that there is no switch in regimes against the alternative of one switch, the appropriate likelihood ratio is

$$\lambda = \frac{1}{2}m\log\hat{\sigma}_1^2 + \frac{1}{2}(T-m)\log\hat{\sigma}_2^2 - \frac{1}{2}T\log\hat{\sigma}^2$$

The estimate of the point at which the switch from one relationship to another has occurred is then the value of m at which λ attains its minimum. However, implementation of this procedures has been hindered by the lack of a distribution theory. Quandt noted on the basis of a Monte

Carlo experiment that a proposed chi-squared approximation to the significant level of the likelihood ratio test is very poor. It was shown empirically in Quandt (1960) that the χ^2 distribution is a poor approximation to that of $-2\log\lambda$.

2.1.2.2. Kim and Siegmund Test

Kim and Siegmund (1989) examined likelihood ratio tests to detect a structural change in simple linear regression when the alternative, H_1, specifies that only the intercept changes and when the alternative, H_2, permits the intercept and the slope to change.

They show that the generalized likelihood ratio test of H_o against H_1 rejected H_o for large values of

$$\max_{T_o \le t \le T_1} \frac{|U_T(t)|}{\hat{\sigma}}$$

where

$$U_T(t) = \left(\frac{t}{1 - t/T}\right)^{\frac{1}{2}} [\bar{y}_t - \bar{y}_T - \hat{\beta}(\bar{x}_t - \bar{x}_T)] \cdot K(x)$$

$$= (\hat{\alpha}_1 - \hat{\alpha}_2)\left(\frac{t}{1 - t/T}\right)^{\frac{1}{2}} \cdot K(x)$$

with

$$K(x) = \left[\frac{1 - t(\bar{x}_t - \bar{x}_T)^2}{(1 - \frac{t}{T})\sum_{k=1}^{t}(x_k - \bar{x}_t)^2}\right]^{-\frac{1}{2}}$$

If it were known that the only possible value of the change

point m is $m=t$, the appropriate two-sample statistic for testing H_o would be $|U_T(t)|/\hat{\sigma}$. The maximization of $|U_T(t)|/\hat{\sigma}$ over t searches for the unknown value of m.

They derived the asymptotic distribution of the test statistic by the following procedure which is the same as that in section 2.1.1.

1) The sequence of test statistic can be written as $|U_T(Tr)|/\hat{\sigma}$.

2) The sequence converges weakly on $[r]$ to $[r(1-r)]^{-1/2}W^o(r)$ by the functional central limit theorem where $W^o(r)$ is the Brownian bridge.

3) Kim and Siegmund choose max function as a mapping from $D[0,1]$ to R. The asymptotic distribution is

$$p(\max_{T_o\le t\le T_1} \frac{|U_T(t)|}{\hat{\sigma}} \ge b) \sim (\frac{2}{\pi})^{\frac{1}{2}} b(1-\frac{b^2}{T})^{\frac{(T-5)}{2}} \int_{t_o}^{t_1} \mu(t)\,\nu[(\frac{2c^2\mu(t)}{1-c^2})^{\frac{1}{2}}]\,dt$$

where

$$\nu(x) = 2x^{-2}\exp\{-2\sum_{n=1}^{\infty} n^{-1}\Phi(-\frac{1}{2}x\sqrt{n})\} \qquad (x>0),$$

$$\mu(x) = \frac{1}{2t(1-t)[1-g^2(t)t(1-t)]},$$

$$g(t) = \frac{\int_0^1 f(u)\,du - t^{-1}\int_0^1 f(u)\,du}{(1-t)[\int_0^1 f^2(u)\,du - (\int_0^1 f(u)\,du)^2]^{\frac{1}{2}}}.$$

and Φ denotes the standard normal distribution function.

4) Instead of using the hitting probability of Brownian bridge, Kim and Siegmund approximated the special function $\nu(x)$ as

$$\nu(x) \approx \exp(-cx)$$

where c is a numerical constant approximately equal to 0.583.

Similarly the likelihood ratio for the test of H_o against H_2 is

$$\max_{t_o \le t \le t_1} \frac{U_{1,T}^2(t) + U_{2,T}^2(t)}{\hat{\sigma}^2}$$

where

$$U_{i,T}(t) = \frac{A_i'(X_{it}'X_{it})^{-1}X_{it}'Y}{\sqrt{A_i'(X_{it}'X_{it})^{-1}A_i)}}$$

with

$$A_1' = (1,-1,0), \qquad A_2' = (0,1,0,-1)$$

$$X_{1t} = \begin{bmatrix} 1 & 0 & x_1 \\ . & . & . \\ . & . & . \\ 1 & 0 & x_t \\ 0 & 1 & x_{t+1} \\ . & . & . \\ . & . & . \\ 0 & 1 & x_T \end{bmatrix}, \qquad X_{2t} = \begin{bmatrix} 1 & x_1 & 0 & 0 \\ . & . & . & . \\ . & . & . & . \\ 1 & x_t & 0 & 0 \\ 0 & 0 & 1 & x_{t+1} \\ . & . & . & . \\ . & . & . & . \\ 0 & 0 & 1 & x_m \end{bmatrix}$$

1) The sequence of test statistic is $\hat{\sigma}^{-1}[U_{1T}(Tr), U_{2T}(Tr)]$.

2) The sequence converges weakly on $[r_1, r_2]$ to

$$[r(1-r)]^{-1/2}[W_1^o(r), W_2^o(r)]$$

by the functional central limit theorem.

3) The asymptotic distribution of the maximum of the sequence is by the continuous mapping theorem

$$(2\pi)^{-1}b^2\left(1 - \frac{b^2}{T}\right)^{\frac{T-6}{2}} \int_{t_o}^{t_1}\int_0^{2\pi} \mu(t,\theta)\,\nu\left[\left(\frac{2c^2\mu(t,\theta)}{1-c^2}\right)^{\frac{1}{2}}\right] d\theta dt$$

4) For calculating the critical values of the limiting

distribution, $v(x)$ is approximated as mentioned above.

Table 1 shows that the critical values for both test statistics by a 10,000 repetition Monte Carlo experiment.

Table 1.
Critical Values of Kim-Siegmund test statistic

T	Prob.	vs. H_1	vs. H_2
20	.10	2.66	2.96
	.05	2.84	3.14
	.01	3.19	3.44
40	.10	2.76	3.12
	.05	2.98	3.33
	.01	3.43	3.73

2.1.2.3. Chu Test

Chu (1989) derived the asymptotic distribution of Quandt's LR statistic in a multiple regression model using the same procedure described in section 2.1.1.

1) The square root of the sequence of the LR statistic is

$$\frac{(Tr)}{\sqrt{T}}\hat{D}_T^{-1/2}(\beta_{(Tr)} - \beta_T)$$

where

$$\hat{D}_T = \left(\frac{\sum x_t' x_t}{T}\right)^{-1}\hat{V}_T\left(\frac{\sum x_t' x_t}{T}\right)^{-1}$$

and $\hat{V}_T$ is a consistent estimator of $V = \lim E[S_T S_T'] / T$ and $S_T = \sum U_T$.

2) By FCLT the square root of the sequence converges weakly to $W^o(r)$ that is a p-dimensional Brownian bridge.

3) By the continuous mapping theorem the maximum of the square root of the sequence converges to $Sup_r \| W^o(r) \|$, so the maximum LR test statistic converges

$$\underset{p \le k \le t-p}{Max}\ 2\ln\Lambda(k) \overset{d}{\to} \underset{r\in(0,1)}{Sup} \frac{[W^o(r)' W^o(r)]}{\sqrt{r(1-r)}}$$

4) The critical values of the limiting distribution can be obtained by the hitting probability of Brownian bridge.

Chu proposed two test statistics for testing constancy of a trend coefficient

$$T_3 \equiv \max\ (\frac{T^{3/2}}{6\sigma})(\frac{m}{T})^3(\beta_m - \beta_T)$$

and co-integration factor

$$\tilde{T}_4 \equiv \max_{m \le T-1} \frac{h}{3\sigma_1} T^{3/2}(\frac{m}{T})^3(\hat{\alpha}_m - \hat{\alpha}_T)$$

where h is a drift parameter. Applying the same procedure described in section 2.1.1 he derived the limiting distributions under the null hypothesis of constant parameters as follows:

$$\lim Pr[|\tilde{T}_3| > c] = \lim Pr[|\tilde{T}_4| > c]$$

$$= Pr[\underset{r\in[0,1]}{Sup} |W^o(t)| > \sqrt{3}\ c$$

Also the critical values can be obtained by procedure 4) in section 2.1.1.

2.1.3. General Alternative Hypothesis

Until now I considered the alternative hypothesis that a single structural change occurred at some unknown point. Here I will discuss the testing procedure of parameter constancy against general unspecified alternatives. An influential paper in this approach is Brown, Durbin and Evans (1975), which proposed the CUSUM and CUSUM of squares tests based on the recursive residuals. However, the CUSUM test turned out to have no asymptotic local power against movements in coefficients of zero-mean regressors. For the power problem of the CUSUM test Ploberger, Kramer and Kontrus (1989) suggested the Fluctuation test and Nyblom (1989) proposed the Locally Most Powerful test.

2.1.3.1. CUSUM Test

The CUSUM test involves considering the plot of the quantity,

$$W_m = \frac{1}{\hat{\sigma}} \sum_{t=k+1}^{m} w_t, \qquad m = k+1, \ldots, T$$

where w_t is the recursive residual. Under H_o, probabilistic bounds for the path of W_m can be determined and H_o is rejected if W_m crosses the boundary (associated with the level of the test) for some m. This test is aimed mainly at detecting systematic movements of coefficients. Against

haphazard rather than systematic types of movements, Brown et al. proposed the CUSUM of Squares test, which uses the squared recursive residuals and is based on a plot of the quantities,

$$S_m = \frac{(\sum_{t=1}^{m} w_t^2)}{S^2} \qquad where \;\; S^2 = \sum_{t=k+1}^{T} w_t^2, \quad m = k+1, \ldots, T$$

The H_o is rejected if the path S_m crossed a boundary determined by the level of the test. These tests are of the goodness-of-fit type in the sense that they seem applicable against a wide variety of alternatives.

Ploberger and Kramer (1986) criticized the idea that the power of the CUSUM test is obtained solely through leverage on the mean of the dependent variable. Thus the CUSUM test has no asymptotic local power against movements in coefficients of zero-mean regressors. The CUSUM of Squares test fares even worse with local asymptotic power equal to size.

2.1.3.2. Fluctuation Test

For the power problem of the CUSUM test, Ploberger, Kramer and Kontrus (1989) proposed the Fluctuation test based on successive parameter estimates rather than on recursive estimates. A similar procedure for single regression model has been suggested by Sen (1980) and the

Fluctuation test was first suggested by Ploberger (1983).

Ploberger, Kramer and Kontrus considered the varying parameter model and proposed the fluctuation test which is based on rejecting the null hypothesis of parameter constancy whenever these estimates fluctuate too much. Their test statistic is

$$S^{(T)} = \max_{t=k,\ldots,T} \frac{t}{\hat{\sigma} T} \|(X^{(T)\prime}X^{(T)})^{1/2}(\beta_i^{(t)} - \beta_i^{(T)}\|_\infty$$

where

$$\hat{\sigma} = [\sum (y_t - x_t^{\prime}\beta^{(T)})^2/(T-K)]^{1/2}$$

and $\|\cdot\|_\infty$ denotes the maximum norm. They derived the limiting distribution of the test statistics using the same procedure described in section 2.1.1.

1) The sequence of the test statistic is written as function of r

$$B^{(T)}(r) = \frac{\tau(r)}{\hat{\sigma} T}(X^{(T)\prime}X^{(T)})^{(1/2)}(\beta^{\tau(r)} - \beta^{(T)})$$

2) By FCLT the sequence, $B^{(T)}(r)$ converges weakly to $W^{o}(r)$ that is a Brownian bridge.

3) By the continuous mapping theorem the test statistic is

$$S^{(T)} = \sup_{0\le r\le 1}\|B^{(T)}\|_\infty$$

4) The critical values of the limiting distribution of the test statistic can be calculated using well known boundary crossing probabilities of the Brownian bridge in section 2.1.1.

Table 2
Critical values of $S^{(T)}$

k	α=0.10	0.05	0.01
1	1.22	1.36	1.63
2	1.35	1.48	1.73
3	1.42	1.54	1.79
4	1.47	1.59	1.83
5	1.51	1.62	1.86

They also show that the fluctuation test has non-trivial local power irrespective of the particular type of structural change. However, Kontrus and Ploberger's (1984) Monte Carlo results show that neither the CUSUM test nor the fluctuation test dominates the other in small samples.

2.1.3.3. Locally Most Powerful Test

Nyblom (1989) developed the locally most powerful test against a parameter variation in the form of a martingale. The martingale specification has an advantage of covering several types of departure of constancy: for example, a single jump at an unknown time point (change point model) or slow random variation (typically random walk). Under the assumption of no parametric changes the joint density f may be written as

$$f(x_1,\dots x_T;\theta_o) = f_1(x_1;\theta_o)\prod_{k=2}^{T} f_k(x_k|x_1,\dots,x_{k-1};\theta_o)$$

Under the alternative it is assumed that the parameter changes obey a martingale and the increment $\theta_k - \theta_{k-1}$ has the covariance matrix

$$E[(\theta_k - \theta_{k-1})(\theta_k - \theta_{k-1})'] = \tau^2 G_k$$

with G_k a known matrix. Then the density for the observations is

$$f^*(x_1,\dots,x_T;\theta_o,\tau^2) = \int\cdots\int f_1(x_1;t_1)$$

$$\times \prod_{k=2}^{T} f_k(x_k|x_1,\dots,x_{k-1};t_k)\,dH(t_1,\dots,t_n)$$

where H is the joint distribution function of $\theta_1,\dots,\theta_T$. For small τ^2 an approximation to f^* is obtained by finding the Taylor expansion of the integrand at θ_o. By imposing sufficient regularity conditions it is obtained that

$$\frac{f^*(x_1,\dots,x_T;\theta_o,\tau^2)}{f(x_1,\dots,x_n;\theta_o)} = 1 + \frac{\tau^2}{2}[\sum_{j=1}^{T}\sum_{k=1}^{T} d_j'(\sum_{i=1}^{\min(j,k)} G_i)\,d_k$$

$$+ \sum_{k=1}^{T} tr(D_k\sum_{i=1}^{k} G_i)] + o(\tau^2)$$

where $d_k = \partial\log f_k/\partial\theta$ and $D_k = \partial^2\log f_k/(\partial\theta\partial\theta')$. The expression in the brackets serves as the locally most powerful test statistic for the null hypothesis $H_o\colon \tau^2=0$ against the alternatives $H_A\colon \tau^2>0$. Nyblom suggested the test statistic as follows:

$$L = \sum_{j=1}^{T}\sum_{K=1}^{T} d_j'(\sum_{i=1}^{\min(j,k)} G_i)\,d_k = \sum_{j=1}^{T}(\sum_{k=j}^{T} d_k')\,G_j(\sum_{k=j}^{T} d_k) > c$$

instead of the full locally most powerful test which includes the second term in the brackets. The proposed test

statistic L based on cumulative sums of the score function is often asymptotically equivalent to the locally most powerful test statistics. If the observations are independent under H_o, then it is sufficient that the elements of G_k and the variances of the elements of D_k are bounded (independently of k). Thus $var[n^{-2}\sum_k tr(D_k \sum_i G_i)]$ tends to 0 and $n^{-2}L$ has a nondegenerate limiting distribution.

The certain assumptions, especially $T^{-1}\sum_{k=1}^{(Tr)} E_{k-1}(d_k d_k') \Rightarrow rJ$, was required to get the limiting distribution of the test statistic using the same procedure in Section 2.1.1.

1) The sequence, L, can be a function of r, that is, the partial sum of the score function is a function of r.

2) By FCLT the partial sums of the score function converges weakly to a p-variate Wiener process $n^{-1/2}\sum_{k=1}^{Tr} d_k \Rightarrow W(r)$ where $W(r)$ has $E[W(r)]=0$ and $Cov[W(s)W(r)] = \min(s,r)J$.

3) By the continuous mapping theorem

$$n^{-2}L \Rightarrow \int_0^1 [W(1) - W(r)]'G(r)[W(1) - W(r)]dr$$

4) Nyblom suggested the Laplace transform of the limiting distribution function to calculate the critical values.

If the parameter differences are assumed to be identically distributed and entail uniform jump probabilities in the change point model, then an appropriate choice is $G(t)=1/J$. And since the starting point θ_o is

rarely known it must be replaced by an estimator $\hat{\theta}_n$. The same procedure is applied to get the limiting distribution of $n^{-2}\hat{L}$.

1) same as above.

2) By FCLT the partial sums of the score function, now, converges weakly to $W^o(r)$ that is a p-dimensional Brownian bridge.

3) By the continuous mapping theorem

$$n^{-2}\hat{L} \Rightarrow \int_0^1 W^o(r)'J^{-1}W^o(r)\,dr = \sum_{k=1}^{\infty} (\pi k)^{-2}\chi_k^2(p)$$

4) Nyblom calculated the critical values of the limiting distribution by Laplace transform of the density.

Table 3
Asymptotic upper-tail percentage points of LMPT

α	df=1	2	3	4	6
0.10	.347	.607	.841	1.063	1.487
0.05	.461	.748	1.000	1.237	1.686
0.01	.743	1.074	1.359	1.623	2.117

Nyblom noted that for large samples a suitable choice is $J_n(\hat{\theta}_n)$ defined by $J_n(\theta_o) = n^{-1}\sum E_{k-1}[d_k(\theta_o)d_k(\theta_o)']$. For the standard linear regression model the test statistic for the constancy of the regression is

$$\hat{L} = tr[S_T^{-1}\sum_{j=1}^{T}(\sum_{k=j}^{T} e_k x_k)(\sum_{k=j}^{T} e_k x_k')]/\hat{\sigma}^2 > c$$

where $e_k = y_k - \hat{\theta}'x_k, \hat{\sigma}^2 = (T-p)^{-1}\sum e_k^2$, and $S_T = T^{-1}\sum x_k x_k'$.

2.1.3.4. Mean Chow Test

Hansen (1990) proposed the MeanChow test which is the average of the Chow sequence. As mentioned above, test statistics of Quandt, Kim and Siegmund and Chu are the MaxChow type test - the maximum of Chow sequence that was designed to detect a single abrupt structural change over sample periods. Hansen's MeanChow statistic is testing for the null hypothesis of parameter constancy against more general alternatives such as the Fluctuation test and the Locally Most Powerful test. He suggested two statistics, one for stationary regressors and the other for non-stationary regressors. Here I will discuss stationary case only; nonstationary case will be discussed in section 2.2. The standard Chow test statistic for coefficient constancy with r known is given by

$$C(r) = b_r' \hat{V}_b^{-1} b_r$$

where b_r is the OLS estimator of β and $\hat{V}_b$ is a consistent estimate of the variance-covariance matrix of b_r.

To get the limiting distribution, the same procedure in section 2.1.1 is applied.

1) The Chow process is given as the function of r.

2) By FCLT and CMT the Chow process converges weakly to the function of the Brownian bridge,

Recall that Chu (1989) derived the square root of the

$$C(r) \Rightarrow \frac{W_1^o(r)' W_1^o(r)}{r(1-r)}$$

sequence first.

3) Hansen constructed the MeanChow test statistic based on the range K = [0.15,0.85] which was suggested by Andrews (1990). This fact comes from the endpoint problem that we need to evaluate the Chow process at points bounded away from the beginning and end of the sample. This is an important component of the asymptotic theory.

$$MeanChow = \frac{1}{T_2 - T_1} \sum_{j=T_1}^{T_2} C(j/T), \qquad T_1 = .15T, \quad T_2 = .85T$$

The limit distribution of MeanChow is by the continuous mapping theorem

$$MeanChow \Rightarrow \int_K C(r)\, dr / \int_K dr$$

4) Hansen tabulated the asymptotic critical values of the test statistics by Monte Carlo methods.

Table 4
Upper percentage points of the MeanChow distribution
Stationary regressors: K = [.15, .85]

k	α=.10	.05	.01
1	2.15	2.86	4.61
2	3.73	4.63	6.52
3	5.18	6.19	8.56
4	6.46	7.60	10.16
5	7.75	8.93	11.61

2.2 Nonstationary Regressors

Concerning the testing procedure of parameter constancy in the model with stationary regressors, after Chow's (1960) test with a known break point, many econometricians mentioned in section 2.1 have developed the tests of parameter constancy with an estimated break point. About the model with nonstationary regressors Perron (1989) initiated the studies concerning a unit root testing procedure that allows for a structural change in the trend function under the alternative hypothesis. Perron's testing procedure is based on a known break point a priori. Christiano's (1988) criticism was that the date of the break ought not be treated as known but rather should be treated as unknown. Following Christiano's criticism, Banerjee, Lumsdaine and Stock (1990), Zivot and Andrew (1990) and Hansen (1990) developed the testing procedures with an estimated break point. Actually these tests are a kind of joint or mixture testing procedure of unit root against stationarity and parameter constancy.

The main problem of making inference about the model with nonstationary regressors is that the limiting distribution of the test statistic depends upon nuisance parameters. Section 2.2.1 discusses the methods that eliminate the nuisance parameter dependency from the

limiting distribution of a test statistic. Section 2.2.2 reviews several testing procedures.

2.2.1. Nuisance Parameter Free Limiting Distribution

The same procedure described in section 2.1.1 can be applied to get the asymptotic distribution of the test statistic for parameter constancy in the model with nonstationary regressors. However, if we allow the disturbances to be correlated with the errors of the nonstationary regressors and the errors are heterogeneously distributed, then the asymptotic distributions become nonstandard in that they depend on the nuisance parameters.

Consider the regression

$$y_t = \mu + \theta t + Ax_t + u_t, \qquad where \quad x_t = x_{t-1} + v_t$$

The basic idea is that if the nonstationary regressor $\{x_t\}$ is strictly exogenous; that is, when $\{x_t\}$ is driven by a process $\{v_t\}$ which is generated independently from the regression error process $\{u_t\}$, then the nuisance parameter vanishes in the limiting distribution of a test statistic.

Let us discuss a little more about the functional central limit theorem in section 2.1.1. If the functional is

$$f^*(B,M) = \int_0^1 dBM'(\int_0^1 MM')^{-1}$$

then $f^*(B,M) \equiv N(0,\Omega\otimes I_p)$ where Ω is the covariance matrix

of a Brownian motion B. This implies that upon appropriate standardization the conventional asymptotic theory applies to regression with integrated regressors if those are strictly exogenous. If not, however, the coefficients in the model with nonstationary regressors converge weakly to a different functional.

Define a functional

$$f(B,M,E) = (\int_0^1 dBM' + E')\ (\int_0^1 MM')^{-1}$$

where B_1 is a Brownian motion concerning nonstationary regressors and B_2 is for regression error process. The matrix M is a stochastic process of continuous sample paths and can be a function of B_2 where E is a matrix (which may be random) of conformable dimension. Park and Phillips (1988) show that the coefficients of the model with integrated regressors converges weakly to this functional with different scales. The coefficient of nonstochastic regressor, A, converges weakly to this functional with T scale, $T(\tilde{A} - A) \Rightarrow f(\cdot)$ and the constant, μ, converges with $\sqrt{T}$ scale, $\sqrt{T}(\tilde{\mu} - \mu) \Rightarrow f(\cdot)$. The coefficient of trend, θ, converges weakly with $T^{3/2}$ scale, $T^{3/2}(\tilde{\theta} - \theta) \Rightarrow f(\cdot)$. The source of the nuisance parameter dependency is E.

Two approaches are suggested to eliminate the nuisance parameter dependency from the limiting distribution of a test statistic.

2.2.1.1. Augmented Dickey-Fuller (ADF) Approach

It is based on the addition of extra lags of first differences of the data as regressors. The use of the ADF statistics is subject to one qualification: the number of lags of the series required for augmented Dickey-Fuller regression to have the correct sample size and good power properties tends to be rather large. With the ADF procedure, the errors are restricted to the class of ARMA(p,q) processes. When the error sequence satisfies a certain assumption, based on arguments outlined in Said and Dickey (1984), the limiting distributions of the test statistics computed from the ADF regression are free of nuisance parameter dependencies. Banerjee, Dolado, and Galbraith (1990), Zivot and Andrews (1990) and Banerjee, Lumsdaine and Stock (1990) follow this approach without a proof of the efficacy of the ADF approach.

2.2.1.2. Nonparametric Transformation

The nonparametric variants are developed by Phillips (1987). Let us define $\Omega = \lim_{T\to\infty} T^{-1}E(S_T S_T')$, $\Sigma = \lim_{T\to\infty} T^{-1}\sum_{1}^{T} E(w_t w_t')$ and $\Lambda = \lim_{T\to\infty} T^{-1}\sum_{t=2}^{T}\sum_{j=1}^{t-1} E(w_j w_t')$ where $w_t = (u_t', v_t')$ and $S_{1t} = \sum_{1}^{t} W_j$. Park and Phillips (1988) show that the source of nuisance

parameter dependency, E, is a function of $\Delta_{21} = \Sigma_{21} + \Lambda_{21}$ where the subscripts mean the partition of the appropriate matrix. They proved that if $\Omega_{21} = \Delta_{21} = 0$ (strict exogeneity of non-stationary regressors), then the conventional asymptotic theory can be applied, especially the Wald-type test statistic is distributed as a chi-square. If not, they proposed a nonparametric transformation of the test statistic to eliminate the parameter dependency.

$$H(\hat{A}) = G(\hat{A}) - 2T\ tr\ \tilde{\Omega}_1^{-1}(\hat{A} - A)\tilde{\Delta}_{21} + T^2\ tr\ \tilde{\Omega}_1^{-1}\tilde{\Delta}_{21}'(\sum_1^T X_t X_t')^{-1}\tilde{\Delta}_{21}$$

The limiting distributions of the H statistics, by construction, do not depend on the nuisance parameters, Δ_{21}.

Hansen (1990) extended this proposition to the test for a structural break with nonstationary regressors. For applying the conventional asymptotic theory two sources of non-normality, Λ_{21} and Ω_{21}, should be eliminated from the functional. He assumed $\Lambda_{21} = 0$ (weak exogeneity of non-stationary regressors) based on the fact that it will vanish if a sufficient number of lagged values of ΔX_{t-k} are included. Then he proposed two stage methods for the nuisance-parameter-free-limiting distribution of the Chow statistics. First he modified the dependent variable, y_t^+, using the OLS estimates of original regression and then obtained the fully modified estimate of A, denoted $\hat{A}^+$, by regressing y_t^+ on x_t. The Chow statistic based on the fully modified estimators has the chi-square asymptotic distribution.

Schwert's (1989) simulations suggested that nonparametric corrections do not perform well even in large samples. He shows that for some (especially MA) error processes, the nonparametric variants of the Dickey-Fuller tests are characterized by low power and incorrect sizes, in samples as large as 1000, and that the use of the ADF test is to be preferred to the use of nonparametrically corrected Dickey-Fuller test statistics.

2.2.2. Several Tests

Here I will review Perron's (1989) testing procedure for a unit root hypothesis with drift at a known break point and then discuss several tests with an unknown break point. Under the assumption of a single structural change Zivot and Andrews (1990) proposed the min-t test and Banerjee, Lumsdaine and Stock (1990) suggested the recursive and sequential tests. Against the general unspecified structural break in the model with non-stationary regressors rather than a single structural change Hansen (1990) developed the MeanChow test. Empirical applications of these testing procedures will be reviewed in Chapter 4.

2.2.2.1. Perron's Hypothesis

Perron (1989) developed a procedure for testing the null hypothesis that a given series has a unit root with drift and that an exogenous structural break occurs at a known time a priori versus the alternative hypothesis that the series is stationary about a deterministic time trend with an exogenous change in the trend function at a known time. He considered three model as follows:

$$Model(A)\quad y_t = \mu^A + \theta^A DU_t + \beta^A t + d^A D(m)_t + \alpha^A y_{t-1} + e_t^A$$

$$Model(B)\quad y_t = \mu^B + \theta^B DU_t + \beta^B t + \gamma^B DT_t^* + \alpha^B y_{t-1} + e_t^B$$

$$Model(C)\quad y_t = \mu^C + \theta^C DU_t + \beta^C t + \gamma^C DT_t^* + d^C D(m)_t + \alpha^C y_{t-1} + e_t^C$$

where

$$D(m) = 1, \quad if\ t = m+1$$
$$= 0, \quad otherwise$$

$$DU_t = 1, \quad if\ t > m$$
$$= 0, \quad otherwise$$

and

$$DT_t^* = t-m, \quad if\ t > m$$
$$= 0, \quad otherwise$$

He assumed that the innovation series $\{e_t\}$ is the ARMA(p,q) type with the order p and q possibly unknown. The null hypothesis of a unit root imposes the following restrictions:

The crash hypothesis:

$$Model(A): \alpha^A = 1,\ \beta^A = 0,\ \theta^A = 0$$

The breaking slope with no crash:

$$Model(B): \alpha^B = 1,\ \gamma^B = 0,\ \beta^B = 0$$

Both effects:

$$Model(C): \alpha^C = 1,\ \gamma^C = 0,\ \beta^C = 0$$

Under the alternative hypothesis of a "trend stationary" process, we expect $\alpha^A,\ \alpha^B,\ \alpha^C < 1,\ \ \beta^A,\ \beta^B,\ \beta^C = 0$ and $\theta^A,\ \theta^C,\ \gamma^B,\ \gamma^C = 0$. Furthermore, under the alternative hypothesis d^A, d^C and θ^B should be close to zero while under the null they are expected to be significantly different from zero.

Perron derived the asymptotic distribution of $T(\hat{\alpha}^i - 1)$ where $i = A,B,C$ and t ratio $t_{\hat{\alpha}^i}$ under the null hypothesis of a unit root. The asymptotic distributions of the various statistics are not influenced by the introduction of the dummy variables $D(m)_t$ since it is a dummy variable affecting a single period. This implies that only two sets of critical values need be evaluated, one corresponding to model (A) and the other to models (B) and (C). He found that the limiting distributions depend on additional nuisance parameters, apart from r (the break-point ratio), $\sigma^2 = \lim_{T\to\infty} T^{-1}E[S_t^2]$ and $\sigma_e^2 = \lim_{T\to\infty} T^{-1}E[\sum_1^T e_t^2]$ where $S_T = \sum_1^T e_t$. In the case of weakly stationary innovations σ^2 is equal to $2\pi f(0)$ where $f(0)$ is the spectral density of $\{e_t\}$ evaluated at frequency zero and σ_e^2 is the variance of the innovations. When the innovation

sequence $\{e_t\}$ is independent and identically distributed $\sigma^2 = \sigma_e^2$, the limiting distribution are invariant with respect to nuisance parameters, except r. Perron tabulated the critical values of the limiting distribution for given value of r by 5,000 times simulation methods. The critical values under the various models are significantly larger than the standard Dickey-Fuller critical values.

2.2.2.2. Zivot and Andrews' Min-t Test

Zivot and Andrews (1990) questioned Perron's exogeneity assumption of r and instead treat the structural break as an endogenous occurrence. Their null hypothesis for the Perron's three models is an integrated process with drift,

$$y_t = \mu + y_{t-1} + e_t$$

The selection of the break point, m, is the outcome of an estimation procedure designed to fit the series to a certain trend stationary representation. They suggested to choose m that gives the least favorable result for the null hypothesis. That is, m is chosen to minimize the one-sided t-statistics for testing $\alpha^i = 1$, when small values of the statistic lead to rejection of the null. Zivot and Andrews proposed the test statistic

$$t_{\hat{\alpha}^i}(\hat{m}^i_{Inf}) = \underset{m \in D}{Inf}\; t_{\hat{\alpha}^i}(m)$$

where D is a specified closed subset of (0,1).

Zivot and Andrews derived the limiting distributions of this statistic when the disturbances are independent and there are no extra lag terms in the regression equations,

$$\operatorname*{Inf}_{m\in D} t_{\hat{\alpha}^i} \Rightarrow \operatorname*{Inf}_{m\in D} \left[\int_0^1 W^i(m,s)^2 ds\right]^{-1/2} \left[\int_0^1 W^i(m,s)\, dW(s)\right] \qquad as \quad T \to \infty$$

As discussed in section 2.3.1 if the disturbances are allowed to be correlated and heterogeneously distributed, then the limiting distribution depends on the nuisance parameters. They followed the ADF approach as Perron used. The errors are assumed to be restricted to the class of ARMA(p,q) processes and the extra lags of first differences of the series are added as regressors. Then the test statistics computed from the ADF regression equations has the limiting distribution free of nuisance parameter dependencies. They tabulated critical values for the limiting distributions by simulation methods. The integral functions are approximated by functions of sums or partial sums of independent normal random variables.

Table 5
Critical values for
the asymptotic distribution of min-t

Model	α=1%	5%	10%
A	-5.34	-4.80	-4.58
B	-4.93	-4.42	-4.11
C	-5.57	-5.08	-4.82

2.2.2.3. Hansen's MeanChow Test

In section 2.1.3 we discussed Hansen's MeanChow test compared with MaxChow test. Hansen (1990) extended the MeanChow test to the model with the non-stationary regressors. As mentioned in section 2.3.1 Hansen follows the nonparametric transformation approach to get the nuisance parameter free asymptotic distribution of the MeanChow test statistic. He developed a two stage method for estimating the fully modified estimator which served the basis to eliminate the parameter dependency from the limiting distribution.

Consider the multiple regression model

$$y_t = A_{1t}x_{1t} + A_{2t}x_{2t} + e_t$$

where x_{1t} is the set of stationary regressors and x_{2t} is the set of non-stationary regressors. His two stage method for the fully modified estimator is that first estimate $\hat{A}_1$ by OLS, and then modify the dependent variable

$$y_t^+ = y_t - \hat{A}_1 x_{1t} - \Omega_{21}\Omega_{11}^{-1}\eta_t$$

where the submatrix of Ω is defined in section 2.3.1 and $\eta_t = \sum_{j=1}^{M} a(j,M)\,\hat{v}_{2t+j}$ with $a(j,M)$, a sequence of fixed weights. Hansen suggested to use the triangular window for the weights

$$a(j,M) = (1 - \frac{j}{M+1})(\frac{2}{M})$$

The fully modified estimate of A_2, denoted by $\hat{A}_2^+$, is

obtained by regressing y_t^+ on x_{2t}. The natural Chow test of a structural break at a known time $[Tr]$ would be given by the OLS regression

$$y_t^+ = \hat{A}_2^+ x_{2t} + \hat{\Gamma}_{2r} x_{2t} I(t \le [Tr]) + \hat{e}_t$$

and the test statistic

$$Chow_2(r) = \hat{\gamma}_{2r}' \hat{V}_{2r}^{-1} \hat{\gamma}_{2r}$$

where $\hat{\gamma}_{2r}$ is the vector of $\hat{\Gamma}_{2r}$. Hansen derived that the limiting distribution of the Chow statistic for any fixed r follows $\chi^2_{n_3(n_2+1)}$ where n_2 is the number of observations of non-stationary regressor and n_3 is the number of equations in the system.

Hansen chooses the mean which maps the Chow process into the real line on the range [0.15,0.85] which is suggested by Andrews (1989). Similarly as shown in section 2.1.3 the MeanChow test of the model with non-stationary regressors is

$$MeanChow = \frac{1}{T_2 - T_1} \sum_{j=T_1}^{T_2} C(j/T), \qquad T_1 = .15T, \quad T_2 = .85T$$

The limit distribution of MeanChow is by the continuous mapping theorem

$$MeanChow \Rightarrow \int_K C(r)\,dr \Big/ \int_K dr$$

Hansen tabulated the asymptotic critical values of the MeanChow statistics by Monte Carlo method.

2.2.2.4. Banerjee, Lumsdaine and Stock's Test

Banerjee, Lumsdaine and Stock (1990) proposed the recursive tests for a unit root and the sequential tests for changes in coefficients. Their motivation for considering recursive unit root tests is that the process might be well approximated as having a unit root over part of the sample but not over another part. The sequential tests are constructed for testing a unit root against some trend-shift/mean-shift alternatives.

First Banerjee et al. considered the model

$$\textit{Model I}:\ y_t = \mu_o + \mu_1 t + \alpha y_{t-1} + \beta(L)\Delta y_{t-1} + e_t$$

where $\beta(L)$ is a lag polynomial of known order p with the roots of $1-\beta(L)L$ outside the unit circle. Under the null hypothesis $\alpha=1$ and $\mu_1=0$. The t-statistic testing the hypothesis that $\alpha=1$ in Model I, computed over the full sample of T observations, is the standard Dickey-Fuller (1979) t-statistic for testing for a unit root, including a constant and a time trend in the regression. Banerjee et al. developed the asymptotic theory for the recursively computed estimators and t-statistics. They found that the recursive estimation of the nuisance parameters does not affect the asymptotic distribution of the recursive Dickey-Fuller statistic based on the block diagonality of the covariance matrix of the transformed regressors and the limiting

distribution of the modified recursive Sargan-Bhargava statistic. They examined five statistics for recursive tests for unit roots (five different mapping from recursive D-F and S-B processes to real line): 1) the full-sample D-F statistic, t_{DF} 2) the maximal D-F statistics, $t_{DF}^{\max}$ 3) the minimal D-F statistic, $t_{DF}^{\min}$ 4) $t_{DF}^{diff} = t_{DF}^{\max} - t_{DF}^{\min}$ 5) the minimal value of the recursive modified S-B statistic, $R^{\min}$. They tabulated the critical values of each statistic by 2,000 Monte Carlo simulation.

Table 6
Critical values for Recursive Unit Root Tests: 10% (5%)

T	t_{DF}	$t_{DF}^{\max}$	$t_{DF}^{\min}$	t_{DF}^{diff}	$R^{\min}$
100	-3.15	-1.93	-3.88	2.95	.0195
	(-3.45)	(-2.21)	(-4.13)	(3.37)	(.0165)
250	-3.13	-1.88	-3.80	2.98	.0199
	(-3.43)	(-2.14)	(-4.07)	(3.36)	(.0170)
500	-3.13	-1.88	-3.82	3.01	.0198
	(-3.42)	(-2.14)	(-4.10)	(3.45)	(.0173)

Note that critical values of $t_{DF}^{\min}$ are different from those of Zivot and Andrews. They reported Monte Carlo results that no single recursive statistic seemed to provide a reliable test against root shifts alternative.

Banerjee et al. constructed the sequential tests for shift or jump in trend at an unknown point. The model

considered is

$$\textit{Model II}: \; y_t = \mu_o + \mu_1 t(m) + \mu_2 t + \alpha y_{t-1} + \beta(L)\Delta y_{t-1} + w' x_{t-1}(m) + e_t$$

The deterministic regressor $t(m)$ captures the possibility of a shift in trend, $t(m) = t - m \; (t > m)$, or jump in the trend, $t(m) = 1 \; (t > m)$, at period m. The sequential statistics are computed using the full sample, sequentially incrementing the date of the hypothetical break, m. They show that the asymptotic distribution of the sequential t-statistic does not depend on the nuisance parameters because the covariance matrix of transformed regressors is block diagonal. Three sequential statistics are examined: 1) the maximum of the sequential F-statistics, $F_T^{\max}$ testing the hypothesis that $\mu_1 = 0$ 2) the sequential D-F statistics evaluated at the value of m that maximizes F_T and $t_{DF}^{m^*}$ 3) the minimal D-F statistic over all the sequentially computed D-F statistics, $t_{DF}^{\min^*}$. Table 7 and 8 show the asymptotic critical values of the three statistics for the case of the trend-shift and the mean-shift.

They found by Monte Carlo studies that the size of the tests is approximately the level of the tests. However the power of these sequential tests shows that the trend break is detected with high probability (the unit root tests reject with high probability). The full-sample standard D-F statistic fails to detect stationarity around a shifting trend, particularly if the break is in the second half of the sample. This confirmed Perron's (1989) results and

Table 7
Critical values of the sequential statistics
Trend shift: 10% (5%)

T	F_T^{max}	$t_{DF}(m^*)$	t_{DF}^{min*}
100	14.30	-4.20	-4.20
	(16.74)	(-4.51)	(-4.51)
250	12.96	-4.10	-4.11
	(15.69)	(-4.41)	(-4.42)
500	13.20	-4.09	-4.11
	(15.29)	(-4.38)	(-4.38)

Table 8
Critical values of the sequential statistics
Mean shift: 10% (5%)

T	F_T^{max}	$t_{DF}(m^*)$	t_{DF}^{min*}
100	15.91	-4.51	-4.52
	(18.40)	(-4.82)	(-4.83)
250	16.42	-4.49	-4.51
	(18.61)	(-4.75)	(-4.75)
500	16.70	-4.53	-4.55
	(19.03)	(-4.79)	(-4.81)

interpretation; the permanent shift in the deterministic trend is mistaken for a persistent innovation to a stochastic trend.

Note that Banerjee et al.'s critical values of min-t statistic are different from those of Zivot and Andrews (1990). Banerjee et al. derived the asymptotic theory for the recursive D-F statistics and S-B statistics and the sequential D-F statistics, but did not derive the exact form of the asymptotic distribution of a specific mapping from the processes of those statistics to the real line, for example, max, min or mean of those statistics (step 4 in section 2.1.1). This may cause the different critical values from those of others.

CHAPTER 3

THE BAYESIAN APPROACH

We begin our discussion of Bayesian inference for a structural change by deriving marginal posterior distributions of parameters. Throughout this chapter I will assume the flat prior distribution of parameters including an unknown break point m based on that the parameters are equally likely at any observation point. This simple flat-prior Bayesian inference will be discussed in detail later. Section 3.2 derives posterior distributions of parameters under the flat prior. Section 3.3 reports some Monte Carlo results on the detection of a break point that the structural change is easier to detect when the series is explosive.

3.1 Posterior Distributions from Flat Prior

Consider the simple two-phase regression model,

$$y_i = \alpha_1 + \beta_1 X_i + \varepsilon_i, \qquad i=1,\ldots,m$$

$$y_i = \alpha_2 + \beta_2 X_i + \varepsilon_i, \qquad i=m+1,\ldots,T$$

Under the assumption of independent normal errors, the

likelihood function is

$$L = (2\pi\sigma^2)^{-T/2} \exp\left[\frac{-1}{2\sigma^2}\left(\sum_{i=1}^{m} (y_i-\alpha_1-\beta_1 X_i)^2 + \sum_{i=m+1}^{T} (y_i-\alpha_2-\beta_2 X_i)^2\right)\right]$$

In this simple case Holbert (1982) derived the posterior density of the change m under the assumption of vague prior densities as follows:

$$P_o(\alpha_1,\beta_1,\alpha_2,\beta_2) \propto constant, \qquad -\infty<\alpha_i,\beta_i<\infty$$

$$P_o(m) = 1/(m-3), \qquad m=2,\ldots,T-2$$

$$P_o(\sigma^2) \propto 1/\sigma^2, \qquad 0<\sigma^2<\infty$$

Combining the prior and likelihood we obtain the joint posterior density of the parameters

$$P_1(m,\sigma^2,\alpha_1,\beta_1,\alpha_2,\beta_2) \propto (\sigma^2)^{-(T/2)+1}\exp\left[\frac{-1}{2\sigma^2}\left(\sum_{i=1}^{m} (y_i - \alpha_1 - \beta_1 X_i)^2 + \sum_{i=m+1}^{T} (y_i - \alpha_2 - \beta_2 X_i)^2\right)\right]$$

Integration on σ^2, α_1, α_2, β_1, β_2 leads to the posterior density of the change point m, giving

$$P_1(m|data) \propto \left[m(T-m)\sum_{i=1}^{m} (X_i-\bar{X}_{i,m})^2 \sum_{i=m+1}^{T} (X_i-\bar{X}_{m+1,T})^2\right]^{-1/2}$$

$$\left[\sum_{i=1}^{m} (y_i-\hat{y}_{i(i,m)})^2 + \sum_{i=m+1}^{T} (y_i-\hat{y}_{i(m+1,T)})^2\right]^{-(T-4)/2}, m=2,\ldots,T-2$$

where $\bar{X}_{k,l} = (\sum_{i=k}^{l} X_i) / (i-k+1)$ and $\hat{y}_{i(k,l)} = \hat{\alpha}_{(k,l)} + \hat{\beta}_{(k,l)} X_i$. Holbert found that the marginal posterior density of the regression coefficients was a mixture of bivariate t densities where the mixing distribution was the marginal posterior mass function of m.

Let us consider the standard multiple regression model

with a single structural change,

$$y = W\theta + \varepsilon, \qquad where \;\; W = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix}, \quad \theta = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}$$

Suppose that ε_i is independently and normally distributed with a zero mean vector and the unknown variance σ^2. Then the likelihood function for m, θ, and σ^2 is given by

$$L(\theta_1,\theta_2,\sigma^2,m|y,W) \propto (\sigma^2)^{-T/2} \exp[-\frac{1}{2\sigma^2}(y - W\theta)'(y - W\theta)]$$

We assume that the change point m is independent of θ and σ^2, and m is equally likely at any observation point between k and $T-k$. We use the flat prior on θ and σ^2, and thus the joint prior is given by

$$P_o(m,\theta,\sigma^2) \propto 1/\sigma^2$$

Combining with the likelihood function we have the joint posterior density function

$$P_1(\theta,\sigma^2,m|y,W) \propto (\sigma^2)^{-(T/2)+1} \exp[-\frac{1}{2\sigma^2}\{S + (\theta-\hat{\theta})'W'W(\theta - \hat{\theta})\}]$$

Integrating out θ and σ^2 yields the marginal posterior mass function

$$P_1(m|y,W) \propto |W'W|^{-1/2}|S|^{-(T-2k)/2}$$

where $|W'W| = |X_1'X_1||X_2'X_2|$ and $S=S_1+S_2$ with $S_i = (y_i - X_i\hat{\theta}_i)'(y_i - X_i\hat{\theta}_i)$ and $\hat{\theta}_i = (X_iX_i)^{-1}X_i'y_i$. The marginal posterior density function is obtained by

$$P_1(\theta|y,W) \propto \sum_{m=k}^{T-k} P_1(\theta|m,y,W)\, P_1(m|y,W)$$

where $P_1(\theta|m,y,W)$ is the conditional marginal posterior density function of θ given m and it is given by a matrix t density

$$P_1(\theta|m,y,W) \propto [I + (\theta - \hat{\theta})'H(\theta - \hat{\theta})]^{-(\nu+2k-1)/2}$$

where $\nu = T - 2k + 1$ and $H = W'W/S$. The marginal posterior densities of θ and σ^2 are mixtures of matrix t densities and the mixing proportion is the marginal posterior probability mass function of the joint function. The marginal posterior density function of θ_i is

$$P_1(\theta_i|y,W) \propto \sum_{m=k}^{T-k} P_1(\theta_i|m,y,W)\, P_1(m|y,W)$$

where $P_1(\theta_i|m,y,W)$ is the conditional marginal posterior density of θ_i given m, and it is a univariate t distribution

$$P_1(\theta_i|m,y,W) \propto \left[1 + \frac{(\theta_i - \hat{\theta}_i)^2}{h^{ii}}\right]^{-(\nu+1)/2}$$

where h^{ii} is the ith diagonal element of H^{-1}. The marginal posterior density of $P_1(\theta_i|m,y,W)$ is a weighted t distribution with weights being the posterior probability mass function of m.

We are interested in the conditional marginal probability mass function of m given a subset of parameters, $\rho_1 = \rho_2 = 1$. We can partition $\theta = [\theta_1' | \theta_2']$ where $\theta_1' = (\rho_1, \rho_2)$, and H is partitioned to conform with the partitioning of θ. Integrating out σ^2 from the joint posterior density of the parameters yields

$$P_1(\theta_1,\theta_2,m|y,W) \propto [I + (\theta_1 - \hat{\theta}_1)'(H_{11} - H_{12}H_{22}^{-1}H_{21})(\theta_1 - \hat{\theta}_1)$$
$$+ \{(\theta_2 - \hat{\theta}_2) + H_{22}^{-1}H_{21}(\theta_1 - \hat{\theta}_1)\}'H_{22}\{(\theta_2 - \hat{\theta}_2) + H_{22}^{-1}H_{21}(\theta_1 - \hat{\theta}_1)\}]^{-T/2}$$

The conditional marginal probability mass function of m given θ_1 can be obtained by integrating out θ_2

$$P_1(m|\theta_1, y, W) \propto [1 + (\theta_1 - \hat{\theta}_1)'(H_{11} - H_{12}H_{22}^{-1}H_{21})(\theta_1 - \hat{\theta}_1)]^{-(T-k_2)/2}$$

where k_2 is the number of elements in θ_2.

3.2 A Monte Carlo Study

In this section I will report some Monte Carlo results about the detection of an unknown break point. The results show that the structural change is well detected in both stationary and nonstationary cases by the marginal posterior distribution of mass function of m.

Consider the simple AR(1) as follow:

$$y_t = \alpha_1 + \rho_1 y_{t-1} + \varepsilon_t, \qquad t = 1, \ldots, m$$

$$y_t = \alpha_2 + \rho_2 y_{t-1} + \varepsilon_t, \qquad t = m+1, \ldots, T$$

The derivations of marginal and conditional posterior distributions of parameters are given in section 3.2.

The behavior of the marginal posterior mass function of m is evaluated by the simulations fixing the values of two α's at 2 and 4 and ρ's at 0.7, 0.95 and 1 and $0.1T \le m \le 0.9T$ from 100 replications. Posterior distributions of m varied much in each iteration of replications. To get the overall view about the behavior of the distribution we averaged each value on the distributions over 100 replications. Thus the distribution being illustrated in the following figures is not the exact distribution but the average behavior of that in the simulation.

Figure 1 shows that the marginal posterior mass function of m has its peak at the break point when samples are drawn by fixing the break point at $m^*=25$ and the parameters at $\alpha_1=2$, $\alpha_2=4$, $\rho_1=\rho_2=1$ and T = 50. The conditional posterior distribution of m given $\rho_1=\rho_2=1$ turns out to be smoother than the marginal distribution. It means that the numbers of detecting the break point at $m^*=25$ by the conditional distribution are less than those by the marginal distribution.

Figure 2 illustrates the marginal and conditional distributions of m when random samples do not have a break by fixing $\alpha_1=\alpha_2=2$ and $\rho_1=\rho_2=1$. The conditional distribution is flat over the sample period, but the marginal distribution is flat within the sample period and high at the beginning and end of the sample.

For stationary series ($\rho_1=\rho_2=0.7$) Figure 3 and 4 plots the marginal and conditional posterior distributions of m. The marginal distribution has a sharp peak at the break point in Figure 4. The conditional posterior distribution, however, does not have a peak around the break point. The numbers of detecting the break point by the conditional posterior distribution of m given $\rho_1=\rho_2=0.7$ are reduced in stationary case, but those by the marginal posterior distribution are not changed (even increased).

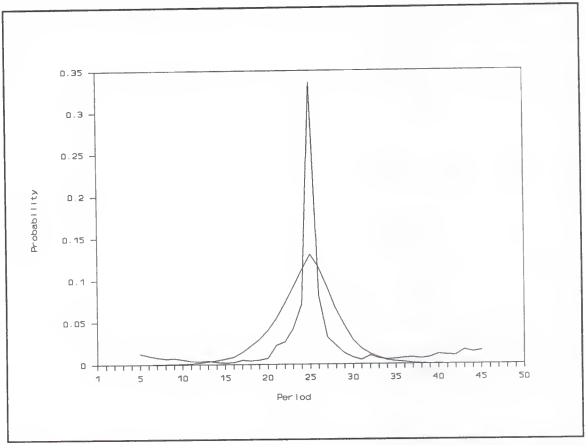

Figure 1
Marginal Posterior Mass Function of m
m^*=25 and rho=1

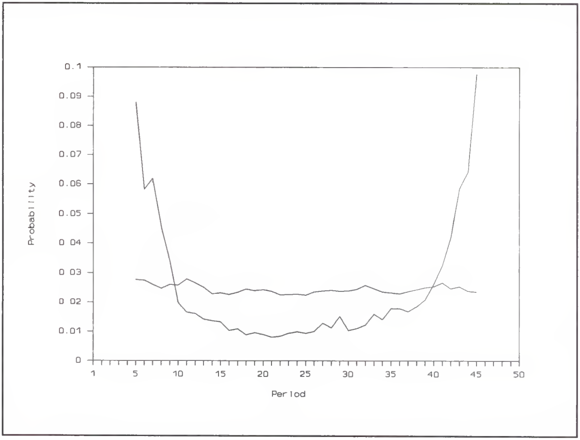

Figure 2
Marginal Posterior Mass Function of m
No Break and rho=1

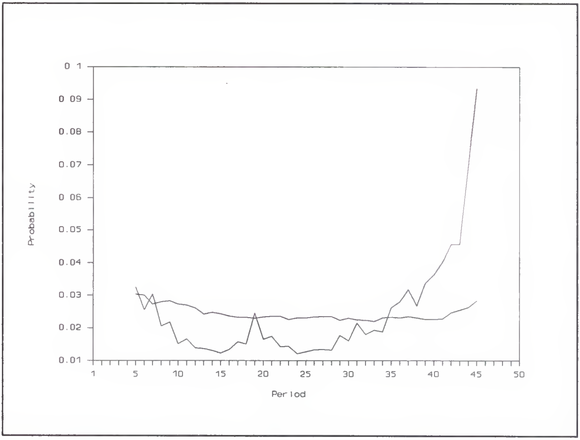

Figure 3
Marginal Posterior Mass Function of m
No Break and rho=0.7

As the sample size becomes bigger, both the marginal and conditional posterior distributions detect well regardless of the stationary and nonstationary case. Figure 4 and 5 illustrate those distributions of nonstationary ($\rho_1=\rho_2=1$; unit root) and stationary case ($\rho_1=\rho_2=0.7$), respectively, when T=100. Note that the peak probability of the marginal distribution in stationary case is higher than that of nonstationary case.

When the autoregressive parameter ρ is changed, from stationary to nonstationary, the marginal posterior distribution of m detected the break point very well. The conditional posterior distribution of m given $\rho_1=0.95(0.7)$ and $\rho_2=1$ appeared to have a flat low probability at period 1 and a high probability at period 2. Figure 6 and 7 show the cases of $\rho_1=0.7$, $\rho_2=1$ and $\rho_1=0.95$, $\rho_2=1$, respectively.

When the series is characterized from nonstationary to stationary, then the marginal posterior distribution detects the break point well. However, asymmetrically the conditional posterior distribution does not show any systematic behavior. Figure 8 and 9 show the reverse case of figure 6 and 7, respectively.

Until now we assumed that a structural break occurred at the middle of the sample period. Figure 10 shows the six cases where the break occurs at 0.15T, 0.2T, 0.4T, 0.6T, 0.8T, and 0.85T with $\alpha_1=2$, $\alpha_2=4$, $\rho_1=\rho_2=1$ and T = 50, respectively. Each peak indicates the assumed

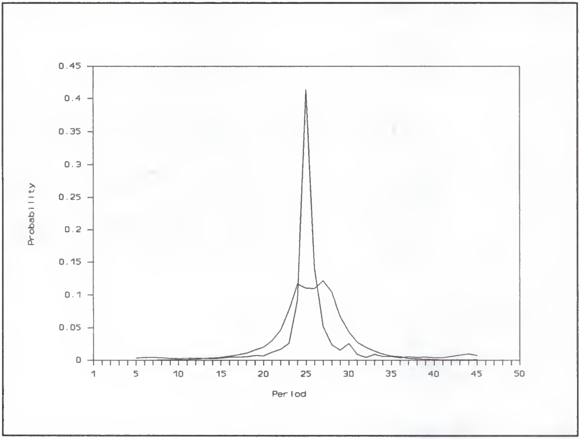

Figure 4
Marginal Posterior Mass Function of m
m^*=25 and rho=0.7

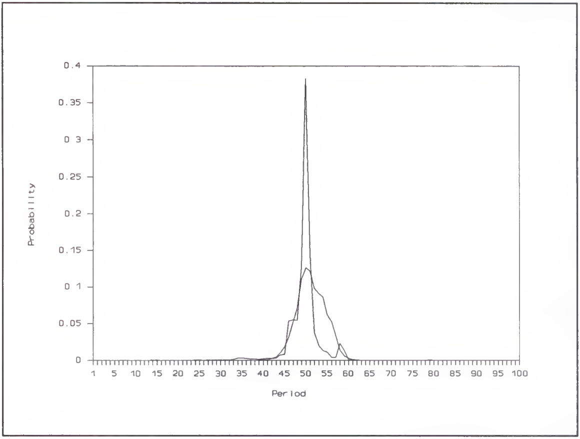

Figure 5
Marginal Posterior Mass Function of m
m^*=50 and rho=0.7

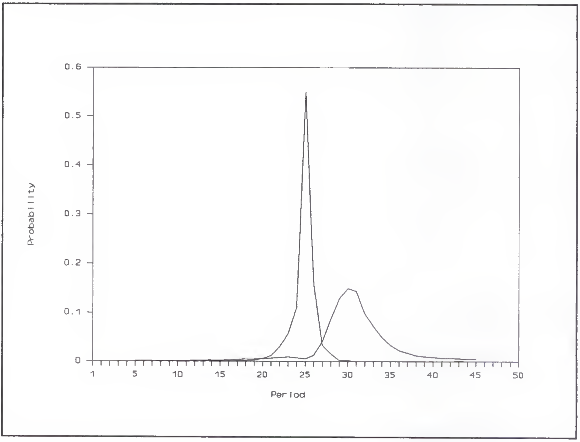

Figure 6
Marginal Posterior Mass Function of m
m^*=25, rho1=0.7 and rho2=1

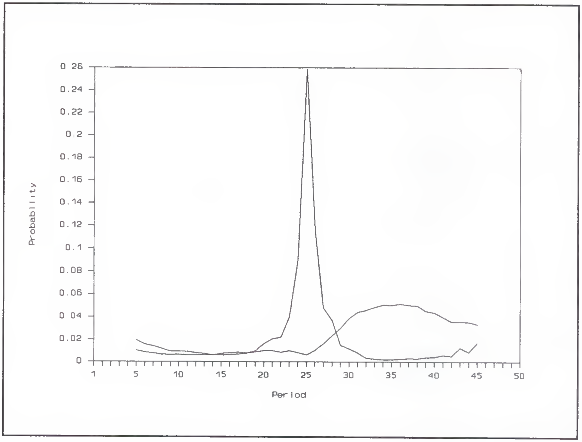

Figure 7
Marginal Posterior Mass Function of m
m^*=25, rho1=0.95 and rho2=1

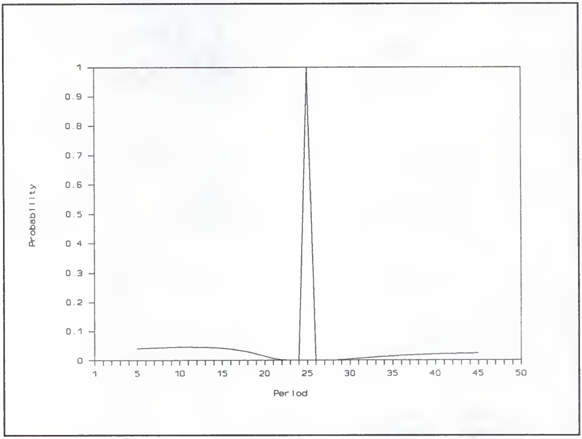

Figure 8
Marginal Posterior Mass Function of m
m^*=25, rho1=1 and rho2=0.7

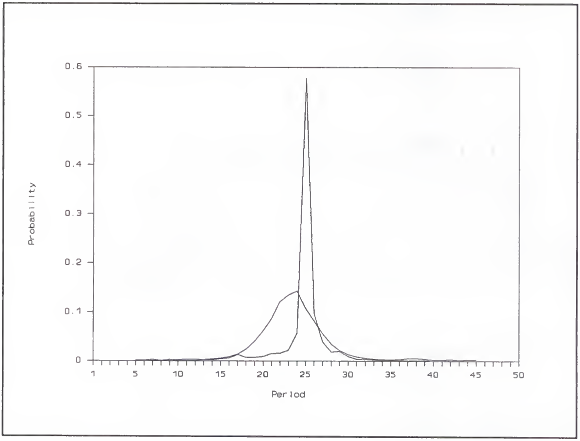

Figure 9
Marginal Posterior Mass Function of m
m^*=25, rho1=1 and rho2=0.95

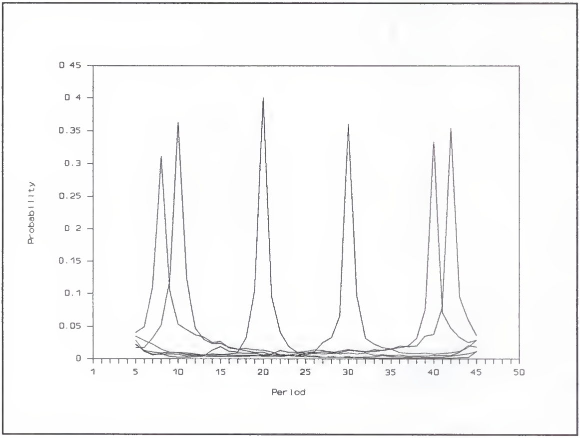

Figure 10
Marginal Posterior Function of m
m^{*}=8,10,20,30,40 and 43 and rho=1

break point exactly. From figure 10 we can infer that the marginal posterior mass function of m detects the structural break sharply whenever a break occurs within the sample period of $0.1T \le m \le 0.9T$.

What will happen if a break occurs at the beginning and end of sample period, $m^*=0.1T$ or $m^*=0.9T$? Figure 11 shows those cases. Peaks are found at the beginning and the end of the sample period as we expected. Comparing figure 11 with figures 2 and 3 indicates that it is hard to infer about the break point when a break occurs at the beginning or the end of the sample period.

In summary the Monte Carlo experiment shows the following:

1) From the average behavior of the marginal posterior distribution of m we can identify a break point as the peak of the marginal posterior distribution of m within the sample period, $0.1T \le m \le 0.9T$.

2) The marginal posterior distribution of m detects a break point well regardless of the stationary and nonstationary case.

3) When the constant is changed, the marginal posterior distribution of m detects the break point sharply, however the conditional posterior distribution of m given ρ's is not as good as the marginal posterior distribution of m for detecting the break point, especially in the stationary case.

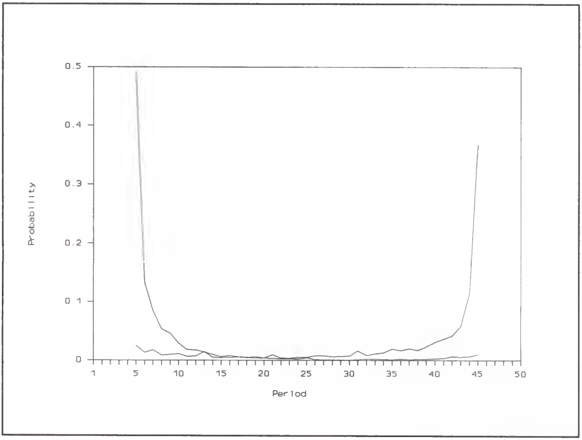

Figure 11
Marginal Posterior Mass Function of m
m^*=5,45 and rho=1

4) When the autoregressive parameter is changed, the marginal posterior distribution of m detects the break point well. The conditional posterior distribution of m shows higher probabilities at nonstationary period than stationary period when the series is changed from stationary to nonstationary, but asymmetrically when the series is changed from nonstationary to stationary, it does not show a systematic behavior.

CHAPTER 4

DATA SETS USED AND PREVIOUS RESULTS

Under the unit root hypothesis random shocks have a permanent effect on the economy and fluctuations are not transitory. A series of empirical analyses followed by Nelson and Plosser (1982) basically confirmed that most macroeconomic variables have a univariate time series structure with a unit root. On the statistical front, besides the standard unit root tests proposed by Dickey and Fuller (1979) and Fuller (1976) there emerged several alternative approaches to test the unit root hypothesis, for example, Phillips and Perron (1986), Campbell and Mankiw (1987, 1988) and Cochran (1986). Empirical applications of these methodologies generally reaffirmed the conclusion that most macroeconomic time series have a unit root.

Recently Perron (1989) proposed modifying the standard Dickey-Fuller test by including dummy variables in the Dickey-Fuller regression in order to allow for a break in the trend and mean. He postulated that the 1929 Great Crash and the 1973 oil price shocks were not a realization of the underlying data-generating mechanism of the various series but could be modelled as exogenous. Then he computed

critical values appropriate to this modified regression and, using the new critical values, found in favor of a structural break in a majority of the time series investigated by Nelson and Plosser (1982). Perron concluded that most macroeconomic time series, 11 out of the 14 series analyzed by Nelson and Plosser, are not characterized by the presence of a unit root and that fluctuations are indeed transitory.

Perron's basic idea is that tests for a unit root using the full sample are biased in favor of accepting the unit root hypothesis if the series has a structural break at some intermediate date. An important criticism of the Perron approach raised by Christiano (1988) was that the break date was assumed to be known. It is more reasonable not to assume a priori knowledge of the break date but rather to allow its estimation to be part of the empirical exercise. This idea initiated the unit root test procedures using sub-samples. Examples include: the recursive test procedures which uses a sequence of statistics constructed by incrementing the sample (Banerjee, Dolado and Galbraith, 1990; Banerjee, Lumsdaine and Stock, 1990; Zivot and Andrew, 1990; Hansen, 1990) and the sequential test procedures which use a sequence of statistics computed by using the full sample at each stage varying dates for the break (Christiano, 1988; Banerjee, Lumsdaine and Stock, 1990). The empirical results by applying these test procedures to the log of U.S. real

GNP provided no evidence against the unit root null in contrast to Perron.

Christiano (1988) criticized Perron's a priori knowledge of a break point and argued that selecting the break date should be a function of the data. He showed that the computed statistical significance level can be drastically different according to the choice of break date. The researcher's break date selection algorithm could then be used to compute the significance level of the test statistic for a break. He suggested the maximal F and minimum significance level techniques as a pre-test examination of the data. Christiano reported that a variety of test statistics reveals no evidence against the null hypothesis of no trend break in post-war Nelson and Plosser's U.S. GNP data using the adjusted critical value by the various break date selection procedures.

Banerjee, Dolado and Galbraith (1990) applied the recursive minimum Dickey-Fuller t-test for unit root null and sequential Dickey-Fuller t-test for the alternative of a trend break proposed by Banerjee, Lumsdaine and Stock (1990) to the Friedman and Schwartz (F-S) (1982) annual series from 1869 to 1975 and Nelson and Plosser (N-P) (1982) annual series from 1909 to 1970. Using the full sample they could reject the null hypothesis of a unit root for the F-S series, but not for the shorter N-P series. For Perron type critics of using full sample in unit root tests, they

computed the recursive minimum D-F t-statistics of two series by using incrementing sub-sample and compared those with the critical values in Banerjee, Lumsdaine and Stock (1990). The recursive test statistics come close to rejecting the unit root null. For the alternative of a trend break against the null of no break in each of the series the sequential statistics offered no significant evidence of a break. These results parallel Christiano's (1988). They also reported that the fact that the normality hypothesis of F-S series' residuals was rejected might cause the application of those critical values hazardous.

Banerjee, Lumsdaine, and Stock (1990) applied several recursive tests and sequential tests to data on postwar real output for seven OECD countries. For OECD data the following description of data sets comes from the Appendix B of Banerjee, Lumsdaine, and Stock (1990).

Data for the United States are GNP from Citibase, for 1947:1 to 1989:2. The data for the six countries come from two sources, the OECD Main Economic Indicators database maintained by Data Resources, Inc. (DRI), and Moore and Moore (1985). In most cases, two series have been spliced together to construct a longer time series of data. Where this has involved an adjustment because the real series are indexed to different base years, they have been adjusted using the earliest available ratio of the two series.

The Canada data are GNP, with 1948:1 to 1960:4 from Moore and Moore and 1961:1 to 1989:2 from DRI. The France data are GDP, 1963:1 to 1989:2, and are from DRI. The French data contain a large negative spike (a strike) in 1968:2; we eliminated this spike by linearly interpolating the value for this quarter. The data for Germany are GNP, with 1950:1 to 1959:4 from Moore and Moore and 1960:1 to 1989:2 from DRI. The data for Italy from DRI were nominal rates, so we have used GDP from Moore and Moore for 1952:1 to 1982:4. The GNP data for Japan is from Moore and Moore for 1952:1 to 1964:4 and from DRI for 1965:1 to 1989:2. The data for the UK are GDP at Factor Cost and are from DRI for 1960:1 to 1989:2.

They first computed the standard Dickey-Fuller statistics and found the hypothesis of one unit root could not be rejected at the 25% level for each of the seven countries. As Perron (1989) emphasized that if the trend-shift/stationary model is correct, then conventional unit root test statistics will incorrectly fail to reject the unit root null. They proposed two types of recursive tests using incrementing sub-sample, the Dickey-Fuller t-statistics and the modified Sargan-Bhargava statistics. For all countries but Italy, the recursive statistics provide no evidence against the unit root null. For Italy the modified Sargan-Bhargava statistics show the evidence against the unit root null, but they reported the poor size performance

of that test statistic. For the alternatives of a trend break Banerjee, Lumsdaine and Stock constructed three different sequential tests for three different alternatives -- trend-shift, mean-break, and a broken drift -- against the unit root null.

Banerjee, Lumsdaine and Stock's results for the USA indicate no rejections of the unit root null against any of the various hypotheses. The sequential test statistics indicate that for Canada, the unit root null is rejected against the mean-shift/stationary alternative, with the break in 1981:3. Thus they interpreted that the recession of the early 1980's is represented as a permanent downward shift in trend growth; after the recovery, output again is stationary along its original growth path. In 1968:2, France experienced a major strike. They used two sets of data, original data and interpolated data over that period, and reported that the results were not changed regardless of the data sets. France seems better characterized as being integrated, but the reduction in the rate of growth of output appears around 1974. For Germany their findings indicated that the unit root null was not rejected, but the constant-drift/unit-root null in favor of the mean-shift/unit-root alternative was less strong. However, this may come from the fact that the earliest observations, near the end of World War II, might have unusually large measurement error. For Italy there is some evidence in favor

of the mean-shift/unit-root alternative around the time of the first oil shock against the constant-drift/unit-root null. For Japan the unit root null is rejected against the trend-shift/stationary alternative, with the break in 1969:4, but on the other hand the broken-drift/unit-root alternative at 1973:2 was favored. Note that Banerjee, Lumsdaine, Stock's results for Japan imply two possible hypotheses; the trend-shift/stationary model with the break in 1969:4 or the broken-drift/unit-root model with the break in 1973:2. For UK the results provide no evidence against the unit root null. They also detected that the growth rate increased in the 80's, although not significantly so.

Zivot and Andrews (1990) proposed the minimum D-F t-test which endogenized the break point selection procedure and reanalyzed the data series considered by Perron using this test statistic. They started with the fact that plots of drifting unit root processes often are very similar to plots of processes that are stationary about a broken trend for some break point. Using their estimated break point asymptotic distributions, they found that there is less evidence against the unit root hypothesis than Perron found for many of the data series. They reversed Perron's conclusions for five of the eleven N-P series for which he rejected the unit root null at the 5% level. Also they reversed Perron's unit root rejection for the postwar quarterly real GNP series from 1947:1 to 1986:3 which come

from the Citibase databank. However, for some of the series (industrial production, nominal GNP, and real GNP), they rejected the unit root hypothesis even after endogenizing the break point selection procedure. For these series, their results provided stronger evidence against the unit root hypothesis than that given by Perron. Note that Zivot and Andrews' (1990) conclusion about the annual N-P GNP series is different from the Banerjee, Dolado and Galbraith's (1990) findings. Recall the different critical values of min-t statistics between Banerjee, Lumsdaine and Stock (1990) and Zivot and Andrews mentioned in Section 2.2.2.

CHAPTER 5

RESULTS FROM THE BAYESIAN PERSPECTIVE

5.1. Introduction

When unit roots are present Bayesian and classical approaches to inference diverge substantially. The asymptotic distribution theory changes discontinuously between the stationary and unit root case. Confidence regions based on asymptotic theory will frequently be disconnected because of the discontinuity in the asymptotic theory. In Bayesian perspectives Sims (1988, p467) argued:

> It has long been recognized that Bayesian inference concerning parameters of linear time series models, conditional on the initial values of the observed sample and Gaussian disturbance distributions, encounters no special difficulties for the case of unit roots. The likelihood, and hence the posterior p.d.f. for a flat prior, is Gaussian in shape regardless of whether or not there are unit (or even explosive) roots. This simple flat-prior Bayesian theory is both a more convenient and a logically sounder starting place for inference than classical hypothesis testing.

We will discuss the debate between Bayesian perspectives and Classical approach to the unit root case and alternative priors instead of flat prior in Chapter 6.

DeJong and Whiteman (1989) adopted a Likelihood Principle to identify the type of prior an investigator

would need to support the inference of integration for the series considered by Nelson and Plosser (1982). The classical Dickey-Fuller test conducted in Nelson and Plosser involved comparing an estimate of AR root, $\hat{\rho}$, to the likelihood function, $l(\hat{\rho}|\rho=1)$. Values of $\hat{\rho}$ in the lower tail constitute evidence against the difference-stationary (DS) hypothesis. Bayesian posterior analysis involves consideration of $l(\rho|\hat{\rho})$ to determine which values of ρ are most likely to have generated the observed data (i.e. generated $\hat{\rho}$). If there is substantial posterior probability associated with values of ρ near unity, the DS inference is supported. They found that this sort of inference was not supported by the data analyzed by Nelson and Plosser and the required prior was excessively sharp. It involves assigning zero probability to the alternative hypothesis that the series are trend-stationary. DeJong and Whiteman infer from their results that evidence in support of a stochastic trend is present for only two series out of 14 series.

In this chapter I will follow the Likelihood Principle approach in Bayesian perspectives and attempt to detect the structural change of the time series used in Nelson and Plosser (1982) and Banerjee, Lumsdaine, and Stock (1990). The simple flat-prior Bayesian method is invariant whether or not there is a unit root in time series. For identifying a structural break we will adopt the Monte Carlo results in Section 3.2 that if there is a break, then a peak in the

marginal posterior distribution of m appears within a sample period.

5.2. Empirical Results

I apply the flat-prior Bayesian analysis of the Chapter 3 to the Nelson-Plosser (N-P) and Friedman-Schwartz (F-S) real per capita GNP series of U.S.A. and quarterly OECD data described in Chapter 4.

For each of nine series we obtain the marginal posterior mass function of m, the marginal posterior distribution of ρ and the conditional posterior distribution functions of ρ given a break point from the augmented Dickey-Fuller regression model with four lags. Table 9 reports the posterior probabilities. Figures in column 2 and 5 to 8 are the posterior probabilities of the near nonstationary set, $P(\rho \geq 0.98)$. The posterior probabilities of Column 2 and 7 to 8 are calculated from the marginal posterior distribution of ρ and Column 5 to 6 from the conditional posterior distribution functions of ρ given a break point. Columns 3 and 4 show the posterior probabilities of a break. Column 3 shows a peak of the posterior mass function of m, when it is found within the sample period, it is a break point. And the posterior

Table 9
Posterior Probabilities

Series	Marg. Dist.	Prob. of Break		Cond. Dist.		Marg. Dist.	
	All	Peak	Prob.	1*	2**	1	2
N-P	.017	beginning	-	-	-	-	-
F-S	.003	end	-	-	-	-	-
USA	.096	1981:3	.169	.055	.001	.128	.001
Canada	.235	1981:2	.391	.017	.006	.028	.102
France	.675	1969:4	.245	.105	.008	.063	.029
Germany	.643	beginning	-	-	-	-	-
Italy	.909	1973:2	.295	.005	.003	.082	.069
Japan	.997	1973:2	.155	.336	.003	.252	.011
UK	.095	end	-	-	-	-	-

* The sample period before the break.
** The sample period after the break.

probability of a break is shown in column 4. When it is found in the beginning or end of the sample period, it is not regarded as a break point.

From the posterior probabilities in Column 2 we can infer that evidence in support of a stochastic trend is present for five OECD countries but not for two countries - USA and UK. For N-P series the result is consistent with DeJong and Whiteman (1989). Banerjee, Dolado, and Galbraith (1990) found the evidence of rejection of a unit root for F-S series. For OECD countries except USA and UK our results parallel Banerjee, Lumsdaine, and Stock's (1990) where their standard Dickey-Fuller test results show a strong evidence of a stochastic trend for all seven countries.

As is clear from the inspection of Column 3 and 4, USA and Canada seem to have a structural change in the recession of the early 1980's (second oil price shock). The 1973 oil price shocks seems to cause a structural change in Italy and Japan. For France the 1968 major strike might be a reason for a break in the output. On the identification of a break our results for Canada and Japan are similar to those of Banerjee, Lumsdaine and Stock. They, however, detected the several different break points for one country according to the different alternative hypothesis.

An inspection from Column 5 to 8 shows that the posterior probabilities for Canada, France, and Italy indicate some evidence supporting Perron's hypothesis. The

outputs of those countries look like nonstationary series, but our results provide evidence for trend-shift/stationary. For Japan before the 1973 oil price shock the output seems to be characterized as being integrated, but after the shock it is better characterized as trend-stationary. The real per capita GNP for USA seems to be a trend-stationary series and has changed the trend around the early 1980's recession (second oil price shock). The output for Germany seems better characterized as being integrated with no break during the sample period. The results for UK provides no evidence for a stochastic trend. Only for Canada are our results parallel to those of Banerjee, Lumsdaine, and Stock. Finally our results for Japan provide an answer for their ambiguous conclusion.

5.3. Probability of a Break and Model Specification

In this section I will show that the posterior distributions of an unknown break point, m, appear to differ according to the different model specifications. In previous sections of this chapter I used the augmented Dickey-Fuller (ADF) regression model for comparing with the results of classical approach as follows:

$$y_t = \alpha + \beta t + \rho y_{t-1} + d(L)\Delta y_{t-1} + \varepsilon_t$$

where four lags of $d(L)$ were included.

When I specified models in different ways, I obtained the different marginal posterior distributions of a break point. Let us consider the simple AR(1) model which excludes the trend term and lags of Δy_{t-1} from the augmented Dickey-Fuller model and the AR(1) with trend when trend term is added to AR(1).

Table 10 shows the break point detected by the three different models. Columns 2, 3 and 4 indicate the results from AR(1), AR(1) with trend, and ADF(4), respectively. For USA the first row of Table 1 shows that no break is detected by the AR(1) model, but when the trend is added to AR(1) model, it gives some evidence of a break at 1981:3. Comparing this break with the ADF(4) case the marginal posterior distribution of m is not the same but the highest probability occurs at the same period.

Similarly for Canada in THE AR(1) case there seems no break, very low probability less than 0.1, but both AR(1) with trend model and ADF(4) model indicate a very sharp peak at 1981:2. For France the AR(1) model seems not to provide strong evidence for a break, but both the AR(1) with trend and the ADF(4) models provide evidence of the same break at 1969:4. The marginal posterior distributions of m of AR(1) and ADF(4) models for Germany indicate no high peak except at the beginning or end of the sample period, that from the

Table 10
Posterior Probabilities of a Break

Series	AR(1)		AR(1) Trend		ADF(4)	
	Peak	Prob.	Peak	Prob.	Peak	Prob.
USA	end	-	1981:3	.162	1981:3	.169
Canada	1961:1	.081	1981:2	.488	1981:2	.391
France	1967:4	.100	1969:4	.424	1969:4	.245
Germany	end	-	1957:1	.136	beginning	-
Italy	end	-	1974:2	.129	1973:2	.295
Japan	1959:1	.197	1959:1	.353	1973:2	.155
UK	end	-	1979:2	.423	end	-

AR(1) with trend model show mild evidence for a break at 1957:1. For Italy the AR(1) model implies no break, but the AR(1) with trend model indicates a break point at 1974:2, though not strongly. Including the lags of series into the model provides strong evidence for a break at the time of the 1973 oil price shock. The AR(1) model and the AR(1) with trend model for Japan provide some evidence for a break at 1959:1, but after adding the lags of series it turns out that the highest probability occurred at 1973:2. For UK, the AR(1) and ADF(4) model show no break, but the AR(1) with trend model indicates relatively strong evidence for a break at 1979:2.

In summary we found the following:

1) The simple AR(1) model indicates no evidence for a break for all countries except Japan.

2) The AR(1) with trend model provide some evidence for a break for all countries.

3) None of the OECD seven countries has the same break by the three different models.

CHAPTER 6

ALTERNATIVE PRIORS FOR THE ANALYSIS OF THE AUTOREGRESSIVE MODEL

In this chapter we will discuss the criticism for the classical unit root inference by Sims (1988) and the alternative priors for Bayesian methodology by Phillips (1990). Sims' basic criticism about the classical inference for unit root is that because the asymptotic distribution theory changes discontinuously between the stationary and unit root cases, classical hypothesis testing based on the asymptotic theory cannot deliver reasonable procedures for inference based on the asymptotic theory. Phillips argued that when the ignorance prior is used instead of the flat prior, the posterior distribution turns out to have a bimodal shape, thus Bayes confidence sets would be disjoint and are therefore formally analogous to those that are generated by classical methods.

Basically Sims argued that the simple flat-prior Bayesian theory is both a more convenient and a logically sounder starting place for inference than classical hypothesis testing. Phillips, however, emphasizes the need to develop a new asymptotic theory for a unit root because his alternative ignorance prior provided the same problem even in Bayesian inferences.

Here we will review Phillips' ignorance prior and provide some Monte Carlo results for comparing the two priors. Phillips argued that since the true value of ρ influences the autocorrelation structure of the time series, flat priors neglect the generic information that is the anticipated amount of information carried by the data about ρ. Strictly speaking this is a property of the likelihood function and the prior should have nothing to do with this argument. Phillips suggested what he calls an objective ignorance prior as follows:

$$P_o \propto |I_{\theta\theta}|^{1/2}$$

where I is the information matrix. This is the prior derived by Jeffreys (1961) using the invariance principle[1]. Consider the simple AR(1)

$$y_t = \rho y_{t-1} + \varepsilon_t \,, \qquad \varepsilon_t \sim iid \;\; N(0,\sigma^2)$$

The flat prior is

$$P_o \propto 1/\sigma$$

and the ignorance prior can be expressed in Phillips' version as follows:

$$P_o \propto (1/\sigma)\, I_{\rho\rho}^{1/2}$$

---

[1]. See also Zellner (1971) p 47.

where

$$I_{\rho\rho} = \frac{T}{1-\rho^2} - \frac{1}{1-\rho^2}\frac{1-\rho^{2T}}{1-\rho^2} + \left(\frac{y_o}{\sigma}\right)^2 \frac{1-\rho^{2T}}{1-\rho^2} \ , \ \rho \neq 1$$

$$= \frac{T(T-1)}{2} + T\left(\frac{y_o}{\sigma}\right)^2 \ , \ \rho = 1$$

To obtain the posterior distribution we multiply the prior with the likelihood function. But the likelihood function already has all the information about data. Thus the posterior distribution from the ignorance prior suggested by Phillips might have a bias because of an exaggerated importance given to ρ based on the structure of the time series.

Phillips' ignorance prior upweights large values of ρ based on the anticipated asymptotic volume of confidence regions which will be tighter when $|\rho| \geq 1$. Sims (1988, p469) argued:

> In the simplest autoregressive model we expect the standard error of estimate of the OLS estimator $\hat{\rho}$ in a sample of given size to be smaller the closer is ρ to 1 (because the sum of squared lagged y's will tend to be larger relative to σ^2 for larger ρ's). This by itself makes it more likely that a given observed $\hat{\rho}$ is a spuriously high estimate generated by a smaller ρ than that it is a spuriously low estimate generated by a larger ρ. But this does not skew the likelihood toward lower ρ's because the distribution of $\hat{\rho}$ is itself skewed to the left for ρ's near 1, which by itself would make it more likely that a given observed $\hat{\rho}$ is spuriously low than it is spuriously high. The classical theory focuses entirely on this latter effect, paying no attention to the danger that we can be misled into giving too much credence to large ρ values because of the more erratic behavior of estimates from models with lower ρ values.

It seems that Phillips suggested his ignorance prior to match the conclusions obtained from the Bayesian approach to those derived from the classical approach when unit roots are present. The likelihood is not skewed toward lower ρ's, thus the only source of skewness is the prior distribution. Phillips' ignorance prior compensates well this skewness to the left for ρ's.

I did a small Monte Carlo experiment for the case T = 50 from 20,000 replications as Phillips (1990) did. The marginal posterior distribution of ρ varied quite q bit between the replications. Figure 12 shows the typical shapes of those distributions when $\rho=1$. However, sometimes the marginal posterior distribution from the ignorance prior turns out to be bimodal as figure 13 shows. One of Phillips' main argument using the ignorance prior is about the disjoint confidence sets in the case of unit root. He argued that because of the bimodality of the marginal posterior distribution function using the ignorance prior, Bayes confidence sets would be disjoint and are therefore formally analogous to those that are generated by classical methods. I did the same Monte Carlo experiment in the case of $\rho=0.5$. Figure 14 indicates the typical types of those distributions. Even in the stationary case the marginal posterior distribution from the ignorance prior turns out to be bimodal in most of the replications out of 20000. Sometimes its shape is as in figure 15. Figures 14 and 15

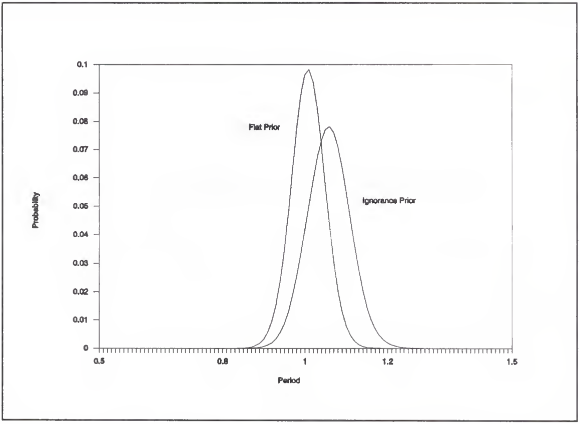

Figure 12
Marginal Posterior Distribution of rho
rho=1

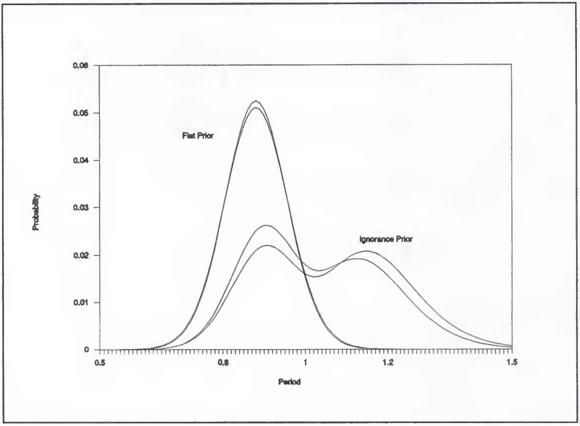

Figure 13
Marginal Posterior Distribution of rho
rho=1

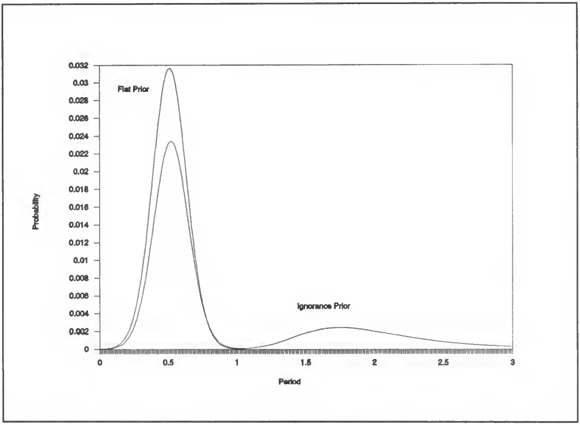

Figure 14
Marginal Posterior Distribution of rho
rho=0.5

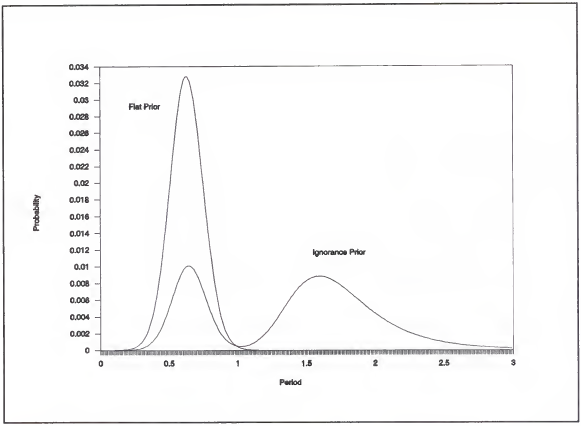

Figure 15
Marginal Posterior Distribution of rho
rho=0.5

imply that the ignorance prior gives too much weights on the nonstationary case even when the true value of ρ is 0.5.

In the normal linear regression model the flat priors lead to Bayesian confidence sets that are equivalent to those in the corresponding sampling theory. Phillips (1990, p8) used ignorance priors in place of flat prior and then argued

> ... Bayesian posteriors for the autoregressive coefficient are frequently bimodal and lead to disjoint confidence sets, just as those based on classical sampling theory asymptotics.

As we saw above, even in the case of stationarity, Phillips' ignorance prior produced the bimodal posterior distributions which are against those in the corresponding sampling theory. Ignorance priors do not, in the case of stationarity, represent prior information in any meaningful sense.

For evidence on the tightness of confidence regions, Table 11, 12 and 13 show the results of Monte Carlo experiments when $\rho=1$, $\rho=0.9$, and $\rho=0.5$, respectively. The first rows in Table 11, 12, and 13 show that the confidence regions from the flat prior are much tighter than those from the ignorance prior even in the case of a unit root. The second rows of Table 11, 12, and 13 indicate the downward bias of the flat prior and also the upward bias of the ignorance prior. The size of the bias from the ignorance prior, however, seems to be bigger than that of the flat prior. From an inspection of the third row of Table 11 and

Table 11
Simulation Results (Null: Rho = 1.0)

	Flat Prior		Ignorance Prior	
	Expect.	Var.	Expect.	Var.
P(.95≤Rho≤1.05)	0.6872	0.0861	0.6351	0.0794
P(1.0≤Rho)	0.4253	0.0869	0.6457	0.0520
Posterior Mean	0.9669	0.0035	1.0116	0.0023
|Mean - Mode|	0.0025	0.0000	0.0192	0.0012
Range	Max.	Min.	Max.	Min.
P(.95≤Rho≤1.05)	1.0000	0.0005	1.0000	0.0008
P(1.0≤Rho)	1.0000	0.0001	1.0000	0.0512
Posterior Mean	1.0975	0.5304	1.5464	0.5904
|Mean - Mode|	0.0771	0.0000	0.6475	0.0000

Table 12
Simulation Results (Null: Rho = 0.95)

	Flat Prior		Ignorance Prior	
	Expect.	Var.	Expect.	Var.
P(.90≤Rho≤1.0)	0.5396	0.0521	0.3945	0.0304
P(.95≤Rho)	0.4379	0.0808	0.6661	0.0489
P(1.0≤Rho)	0.1692	0.0229	0.4572	0.0317
Posterior Mean	0.9144	0.0051	0.9872	0.0040
|Mean - Mode|	0.0025	0.0000	0.0421	0.0026
Range	Max.	Min.	Max.	Min.
P(.90≤Rho≤1.0)	0.8508	0.0001	0.7103	0.0002
P(.95≤Rho)	1.0000	0.0000	1.0000	0.0091
P(1.0≤Rho)	0.9567	0.0000	0.9970	0.0090
Posterior Mean	1.0625	0.4239	1.4797	0.4392
|Mean - Mode|	0.0800	0.0000	1.2714	0.0000

Table 13
Simulation Results (Null: Rho = 0.50)

	Flat Prior		Ignorance Prior	
	Expect.	Var.	Expect.	Var.
P(.45≤Rho≤0.55)	0.2457	0.0087	0.1997	0.0067
P(.50≤Rho)	0.4875	0.0819	0.5777	0.0752
P(1.0≤Rho)	0.0004	0.0000	0.1617	0.0157
Posterior Mean	0.4818	0.0145	0.7135	0.0525
|Mean - Mode|	0.0041	0.0001	0.2157	0.0251
Range	Max.	Min.	Max.	Min.
P(.45≤Rho≤0.55)	0.3974	0.0001	0.3922	0.0000
P(.50≤Rho)	1.0000	0.0003	1.0000	0.0016
P(1.0≤Rho)	0.1000	0.0000	0.9924	0.0003
Posterior Mean	0.8627	0.0893	2.5054	0.0931
|Mean - Mode|	0.1132	0.0000	1.5061	0.0000

the fourth rows of Table 12 and 13, the posterior means from ignorance priors seem to be less biased that those from flat priors. However, as we discussed above, comparing the posterior means is meaningless because of the bimodality of the marginal posterior distribution of the autoregressive coefficient from the ignorance priors. The last rows of Table 11, 12, and 13 indicate the difference between the posterior means and the posterior modes. These figures imply that the bimodality from the ignorance priors occurs very frequently as ρ deviates from 1.

Sims (1988, p467) suggested a better classical approach:

> ... Using Monte Carlo small sample distribution theory ...One can generate the joint distribution of test statistics of interest for a number of parameter points near the likelihood-maximizing one and compare the likelihood of the observed sample under the various estimated joint distributions. ... doing it systematically eventually leads back to a Bayesian framework.

Sims also argues that the analytical difficulties of the classical inferential procedures even in simple cases prevent our making progress on the real issues - for example, nonnormality of disturbances and proper accounting for the evidence about parameters contained in initial conditions.

CHAPTER 7

CONCLUSION

The two crucial parameters in the models considered are m, the break-point and ρ, the autoregressive parameter. Monte Carlo studies about detecting a structural change in the autoregressive model show that the Bayesian posterior mass function of m detects a break point more readily than the classical approach even when the series is nonstationary. When a peak of the marginal posterior mass function of m occurs within a sample period, it indicates a break point. Bayesian methodology and the results of Monte Carlo experiments are applied to the data sets analyzed by Banerjee, Dolado, and Galbraith (1990) and Banerjee, Lumsdaine, and Stock (1990).

An inspection of the marginal posterior distribution of ρ by using full samples shows evidence for a stochastic trend for five OECD countries, but not for the Nelson-Plosser, and Friedman-Schwartz annual series for the US, and the quarterly series for US and UK. The results for N-P and F-S series is consistent with those obtained by DeJong and

Whiteman (1989) and different from those obtained by the classical approach. For the five OECD countries -- Canada, France, Germany, Italy, and Japan -- except Germany, using the Bayesian inference of the marginal posterior mass function of m, I found strong evidence supporting Perron's hypothesis even after endogenizing the break point selection procedure. The results for four of the seven OECD countries -- Canada, France, Italy, and Japan -- show that standard tests of the unit root hypothesis were biased in favor of accepting the unit root hypothesis if the series had a structural break at some intermediate date. The results for Japan provide some evidence that the time series on output has changed from nonstationary to stationary around the 1973 oil price shock. This may explain Banerjee, Lumsdaine, and Stock's ambiguous conclusions among several alternatives for Japan.

Recently Phillips (1990) criticized the simple flat-prior Bayesian approach and suggested an alternative ignorance prior based on Jeffreys' principle of invariance. The Ignorance-prior Bayesian approach suggested by Phillips, which results in the same inference as that of the classical approach, uses the sample information by giving more weight to large values of ρ. Phillips argued that the ignorance-prior tightens the confidence regions near the true value.

My results of a small Monte Carlo experiment provides evidence against Phillips' argument. My empirical applications of the flat-prior Bayesian approach for a structural break result in divergent conclusions from those of the classical approach.

REFERENCES

Andrews, Donald W. K. (1990), "Tests for Parameter Instability and Structural Change with Unknown Change Point," Cowles Foundation Discussion Paper No. 943.

Billingsley, P. (1968), Convergence of Probability Measures, Wiley, New York.

Banerjee, A., J. Dolado and J. W. Galbraith (1990), "Recursive Tests for Unit Roots and Structural Breaks in Long Annual GNP Series," unpublished manuscript, University of Florida.

Banerjee, A., R. L. Lumsdaine and J. H. Stock (1990), "Recursive and Sequential Tests of the Unit Root and Trend Break Hypotheses: Theory and International Evidence," unpublished manuscript, University of Florida.

Booth, N. B. and A. F. M. Smith (1982), "A Bayesian Approach to Retrospective Identification of Change-points," Journal of Econometrics 19, 7-22.

Broemeling, L. D. and H. Tsurumi (1987), Econometrics and Structural Change, Marcel Dekker, Inc., New York.

Brown, R. L., J. Durbin, and J. M. Evans (1975), "Techniques for Testing the Constancy of Regression Relationships over Time," Journal of the Royal Statistical Society, Series B, 37, 149-192.

Campbell, J. Y. and N. G. Mankiw (1987), " Permanent and Transitory Components in Macroeconomic Fluctuations," American Economic Review: Papers and Proceedings 77, 111-117.

Campbell, J. Y. and N. G. Mankiw (1988), "Are Output Fluctuations Transitory?" Quarterly Journal of Economics 102, 875-880.

Chow, G. C. (1960), "Tests of Equality between Sets of Coefficients in Two Linear Regressions," Econometrica 28, 591-605.

Christiano, L. J. (1988), "Searching for a Break in GNP," NBER Working Paper No. 2695.

Chu, C. J. (1989), "New Tests for Parameter Constancy in Stationary and Nonstationary Regression Models," unpublished manuscript, Department of Economics, University of California at San Diego.

Cochran, J. H. (1988), "How Big is the Random Walk in GNP?" Journal of Political Economy 96, 893-920.

Davies, R. B. (1977), "Hypothesis Testing When a Nuisance Parameter is Present Only under the Alternative," Biometrika 64, 247-254.

DeJong, D. N. (1989), "Co-integration and Trend Stationarity in Macroeconomic Time Series: Evidence from the Likelihood Function," Working Paper No. 89-05, University of Iowa.

DeJong, D. N., J. C. Nankervis, N. E. Savin and C. H. Whiteman (1988), "Integration versus Trend-stationarity in Macroeconomic Time Series, Working Paper No. 88-27a, University of Iowa.

DeJong, D. N. and C. H. Whiteman (1989), "Trends and Random Walks in Macroeconomic Time Series: A Reconsideration Based on the Likelihood Principle," Working Paper No. 89-04, University of Iowa.

Diaz, J. (1982), "Bayesian Detection of a Change of Scale parameter in Sequences of Independent Gamma Random Variables," Journal of Econometrics 19, 23-29.

Dickey, D. A. and W. A. Fuller (1979), "Distribution of the Estimator for Autoregressive Time Series with a Unit Root," Journal of the American Statistical Society 74, 427-431.

Edwards, A. W. F. (1972), Likelihood, Cambridge University Press, Cambridge.

Friedman, M. and A. J. Schwartz (1982), Monetary Trends in the United States and the United Kingdom: Their Relation to Income, Prices and Interest Rates, 1867-1975, University of Chicago Press, Chicago.

Goldfeld, S. M. and R. E. Quandt (1972), Nonlinear Methods in Econometrics, North-Holland, Amsterdam.

Hansen, B. E. (1990), "Testing for Structural Change of Unknown Form in Models with Non-stationary Regressors," unpublished manuscript, Department of Economics, University of Rochester.

Harvey, A. C. (1981), The Econometric Analysis of Time Series, Phillip Allan, Oxford.

Hinkley, D. V. (1969), "Inference about the intersection in two-phase regression," Biometrika 56, 495-504.

Hinkley, D. V. (1971), "Inference in Two-phase Regression," Journal of the American Statistical Society 66, 736-743.

Holbert, D. (1982), "A Bayesian Analysis of a Switching Linear Model," Journal of Econometrics 19, 77-87.

Jeffreys, H. (1961), Theory of Probabilities 3rd ed., Oxford University Press, Oxford.

Kim, H. J. and D. Siegmund (1989), "The Likelihood Ratio Test for a Change-point in Simple Linear Regression," Biometrika 76, 409-423.

Kloek, T. and H. K. van Dijk (1978), "Bayesian Estimates of Equation System Parameters: An Application of Integration by Monte Carlo," Econometrica 46, 1-19.

Kontrus, K. and W. Ploberger (1984), "A Simulation Study of Some Tests for Parameter Constancy in Linear Models," Paper given at the Econometric Society European Meeting, Madrid.

Maddala, G. S. (1977), Econometrics, McGraw-Hill, New York.

Moore, G. H. and M. H. Moore (1985), International Economic Indicators. A Sourcebook. Greenwood Press, Westport, CN.

Nelson, C. R. and C. I. Plosser (1982), "Trends and Random Walks in Macro-economic Time Series: Some Evidence and Implications," Journal of Monetary Economics 10, 139-162.

Nyblom, J. (1989), "Testing for the Constancy of Parameters over Time," Journal of the American Statistical Association 84, 223-230.

Park, J. Y. and P. C. B. Phillips (1988), "Statistical Inference in Regressions with Integrated Processes: Part 1," Econometric Theory 4, 468-497.

Perron, P. (1989), "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis," Econometrica 57, 1361-1401.

Phillips, P. C. B. (1987), "Time Series Regression with a Unit Root," Econometrica 55, 277-301.

Phillips, P. C. B. (1990), "To Criticize the Critics: An Objective Bayesian Analysis of Stochastic Trends," Cowles Foundation Discussion Paper No. 950, Yale University.

Phillips, P. C. B. and S. N. Durlauf (1986), "Multiple Time Series Regression with Integrated Process," Review of Economics Studies 53, 473-495.

Phillips, P. C. B. and P. Perron (1986), "Testing for a Unit Root in Time Series Regression," Cowles Foundation Discussion Paper No. 795, Yale University.

Ploberger, W. (1983), "Testing the Constancy of Parameters in Linear Models," Paper given at the Econometric Society European Meeting, Pisa.

Ploberger, W. and W. Kramer (1986), "The Local Power of the CUSUM and CUSUM of Squares Tests," unpublished manuscript, Technische Universitat, Vienna.

Ploberger, W., W. Kramer, and K. Kontrus (1989), "A New Test for Structural Stability in the Linear Regression Model," Journal of Econometrics 40, 307-318.

Quandt, R. E. (1958), "The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes," Journal of the American Statistical Association 53, 873-880.

Quandt, R. E. (1960), "Tests of the Hypothesis That a Linear Regression System Obeys Two Separate Regimes," Journal of the American Statistical Association 55, 324-330.

Said, S. E. and D. A. Dickey (1984), "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order'" Biometrika 71, 599-608.

Salazar, D. (1982), "Structural Changes in Time Series Models," Journal of Econometrics 19, 147-163.

Sargan, J. D. and A. Bhargava (1983), "Testing Residuals from Least Squares Regression for Being Generated by the Gaussian Random Walk." Econometrica 51, 153-174.

Schotman, P. and H. K. van Dijk (1989), "A Bayesian Analysis of the Unit Root Hypothesis," unpublished manuscript, Erasmus University.

Schwert, G. W. (1989), "Tests for Unit Roots: A Monte Carlo Investigation," Journal of Business and Economic Statistics 7, 147-159.

Sen, P. K. (1980), "Asymptotic Theory of Some Tests for a Possible Change in the Regression Slope Occurring at an Unknown Time Point," Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebeite 52, 203-218.

Sims, C. A. (1988), "Bayesian Skepticism on Unit Root Econometrics," Journal of Economic Dynamics and Control 12, 463-474.

Sims, C. A. and H. Uhlig (1988), "Understanding Unit Roots: A Helicopter Tour," Federal Reserve Bank of Minneapolis Institute for Empirical Macroeconomics, Discussion Paper No. 4.

Stock, J. H. and M. W. Watson (1986), "Does GNP have a Unit Root?" Economics Letters 22, 147-151.

Stock, J. H. and M. W. Watson (1988), "Variable Trends in Economic Time Series," Journal of Economic Perspective 2, 147-174.

van Dijk, H. K. (1987), "Some Advances in Bayesian Estimation Methods Using Monte Carlo Integration," Advances in Econometrics 6, 215-261, JAI Press Inc.

Zellner, A. (1971), Introduction to Bayesian Inference in Econometrics, Wiley, New York.

Zivot, E. and D. W. K. Andrews (1990), "Further Evidence on the Great Crash, the Oil Price Shock and the Unit Root Hypothesis," unpublished manuscript, Cowles Foundation, Yale University.

BIOGRAPHICAL SKETCH

In-Moo Kim was born in Pusan, Korea, on January 6, 1958. In 1981, he received B.A. degree in economics from the Sung Kyun Kwan University in Seoul, Korea. After he graduated in 1984 from the Graduate School of Sung Kyun Kwan University with M.A. degree in economics, he took a position of full-time instructor at Shin-gu Junior College in Seongnam city, Korea, and worked as chairman of the Department of Computer Science at same college from 1985 to 1986. In 1987, he entered the doctoral program in economics at the University of Florida and anticipates completing his doctoral studies in 1991.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

G.S. Maddala, Chairman
Graduate Research Professor of Economics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Mark Rush
Associate Professor of Economics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Prakash Loungani
Assistant Professor of Economics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

John S. Shonkwiler
Professor of Food and Resource Economics

This dissertation was submitted to the Graduate Faculty of the Department of Economics in the College of Business Administration and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

May, 1991

Dean, Graduate School